

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 03:09:29 ; Search time 5708 Seconds  
(without alignments)  
10441.471 Million cell updates/sec

Title: US-10-690-564-1

Perfect score: 1230  
Sequence: 1 ccacgcgcctccgaacactca.....aaaaaaaaaaaaaaaaaaaa 1230

Scoring table: IDENTITY\_NTC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: \*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sbs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1230	100.0	1230	8	AF083913	AF083913 Arabidops
2	1197.4	97.3	1208	8	AY072347	AY072347 Arabidops
3	1162.4	94.5	1192	8	AY086570	AY086570 Arabidops
4	1139	92.6	1139	8	BT003359	BT003359 Arabidops
5	1124.2	91.4	1206	8	ATANNEX	X99224 A.thaliana
6	1120	91.1	1159	8	ATU28415	U28415 Arabidops
7	1016.2	82.6	81513	8	FL4D7	AC021198 Sequence
8	954	77.6	954	6	CQ086050	CQ086050 Sequence
9	954	77.6	954	8	AF332435	AF332435 Arabidops
10	790.8	64.3	954	8	AY356355	AY356355 Brasiica
11	497.8	40.5	1141	8	GHU89609	U89609 Gossypium h
12	488.6	39.7	948	8	GHU73746	U73746 Gossypium h
13	478.6	38.9	1112	8	AF006197	AF006197 Lavatera
14	471.8	38.4	2588	6	BD236006	BD236006 Materials
15	468.6	38.1	1293	6	BD236034	BD236034 Materials
16	461.2	37.5	1180	6	AR231102	AR231102 Sequence
17	458	37.2	1420	8	BT012975	BT012975 Lycopersi
18	457.2	37.2	1160	8	AY085713	AY085713 Arabidops

20	455.6	37.0	1137	8	AF083914	AF083914 Arabidops
21	455.6	37.0	1230	8	AY070400	AY070400 Arabidops
22	453.2	36.8	954	6	CQ086274	CQ086274 Sequence
23	453.2	36.8	985	8	AY096577	AY096577 Arabidops
24	451	36.7	1080	8	CANI30956	AJ130956 Capsicum
25	450.6	36.6	1086	8	AY351650	AY351650 Gossypium
26	449.6	36.6	1157	6	AX927134	AX927134 Sequence
27	449.6	36.6	1197	8	AP113545	AP113545 Nicotiana
28	449.2	36.5	945	8	STU401032	AJ401032 Solanum t
29	447.6	36.4	1160	8	AF079232	AF079232 Lycopersi
30	445.4	36.2	1196	8	NTANNB511	Y14972 Nicotiana t
31	444.8	36.2	1182	8	MSANN	X74947 M.sativa mr
32	436.2	35.5	1172	8	NTANNB671	Y14973 Nicotiana t
33	430.4	35.0	1005	8	AY014799	AY014799 Arabidops
34	427.2	34.7	1242	8	AF079231	AF079231 Lycopersi
35	426.2	34.7	1146	8	GHU73747	U73747 Gossypium h
36	393.8	32.0	1003	8	AY014798	AY014798 Arabidops
37	393.8	32.0	1120	8	AK175572	AK175572 Arabidops
38	393.8	32.0	1120	8	AK175641	AK175641 Arabidops
39	393.8	32.0	1120	8	AK175892	AK175892 Arabidops
40	389	31.6	472	11	BX548120	BX548120 Arabidops
41	364.2	29.6	789	6	BD236004	BD236004 Materials
42	364.2	29.6	789	6	BD236037	BD236037 Materials
43	364.2	29.6	789	6	AR231105	AR231105 Sequence
44	346.6	28.2	1277	8	ZMANNP33	X98244 Z.mays mRNA
45	344	28.0	1223	8	AK101787	AK101787 Oryza sat

## ALIGNMENTS

RESULT 1  
AF083913  
LOCUS AF083913 1230 bp mRNA linear PLN 02-JUN-1999  
DEFINITION Arabidopsis thaliana annexin (AnnA1) mRNA, complete cds.  
ACCESSION AF083913  
VERSION AF083913.1 GI:4959105  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopses.  
1 (bases 1 to 1230)  
Clark,G.B. and Roux,S.J.  
Isolation and Characterization of Two Different Arabidopsis Annexin  
cDNAs (Accession Nos. AF083913 and AF083914) (PCR 99-065)  
JOURNAL  
REFERENCE  
AUTHORS Clark,G.B. and Roux,S.J.  
TITLE Direct Submision  
JOURNAL Submitted (12-AUG-1998) Botany, University of Texas, 24th at  
Whitis, Austin, TX 78713, USA  
FEATURES  
source 1..1230  
Location/Qualifiers  
1..1230  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/cultiivar="Columbia"  
/db\_xref="taxon:3702"  
1..1230  
/gene="AnnA1"  
34..987  
/note="Calcium-binding protein"  
/codon\_start=1  
/product="annexin"  
/protein\_id="AAD34236.1"  
/db\_xref="GI:4959106"  
/translation="MATTAKVSDSPAPDDAQLQTAFEGCTNEDLIITSIARSAE  
ORKVIRQAVHYETGDLTKTDLKESNDPERAIIIMLTLEPERDALLNENATKWTSS  
NOVLMEVETRTSTGLHARQAVHARVYKKSLEEDVAHHTTGGFRLLVLSVSYREG  
DEVNMTLAKQSAKLVHEKIKDGHVDEIVIRLSTRSQAQINATFNRYODDGEELIK  
SLEEGDDDDKFLALIRSTIGCTRELYFVIVLSAINKGTGDEGALRIYVTTAEID

## ORIGIN

LNVIGBYORNSIPLEKAITKDTRGDYERMLVALLGENDA"

Query Match 100.0%; Score 1230; DB 8; Length 1230;  
 Best Local Similarity 100.0%; Pred. No. 8,1e-301;  
 Matches 1230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 CCAAGCGTCGAAACAATAAAGTAGAAGAAAAATGCGACCTTTAAGGTTTCTGATTC 60
DB      1 CCAAGCGTCGAAACAATAAAGTAGAAGAAAAATGCGACCTTTAAGGTTTCTGATTC 60
QY      61 GTTCCCTGCTCTTCTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGGATAC 120
DB      61 GTTCCCTGCTCTTCTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGGATAC 120
QY      121 AACGAGACCTTGATCATATCAATCTTGGCTCACAGAAAGTGTGAACAGAGAAAGTCATC 180
DB      121 AACGAGACCTTGATCATATCAATCTTGGCTCACAGAAAGTGTGAACAGAGAAAGTCATC 180
QY      181 AAGCAAGACATACCAAGAACTTACGCGAAGACCTTCTCAAGACTTGTGACAAGAGCTC 240
DB      181 AAGCAAGACATACCAAGAACTTACGCGAAGACCTTCTCAAGACTTGTGACAAGAGCTC 240
QY      241 TCTTAAGATTTGAGAGAGACTTATCTTGTGTGACTTTGAACCCGCTGAGGTGATGCT 300
DB      241 TCTTAAGATTTGAGAGAGACTTATCTTGTGTGACTTTGAACCCGCTGAGGTGATGCT 300
QY      301 TTATTGGCTAATGAACTCAAAAAGATGSACTTCAAGAACCAAGTTCTTATGGAAGTT 360
DB      301 TTATTGGCTAATGAACTCAAAAAGATGSACTTCAAGAACCAAGTTCTTATGGAAGTT 360
QY      361 GCTTGACAAAGACATCAACGCGAGCTGCTTCAAGCTAGGCAAGCTTACCATCTCGCTAC 420
DB      361 GCTTGACAAAGACATCAACGCGAGCTGCTTCAAGCTAGGCAAGCTTACCATCTCGCTAC 420
QY      421 AAGAACTCTTGTGAAGAGAGCTTGTCAACACACTACCGGTGACTTCAAAAAGCTTTTG 480
DB      421 AAGAACTCTTGTGAAGAGAGCTTGTCAACACACTACCGGTGACTTCAAAAAGCTTTTG 480
QY      481 GTTTCCTGTTACCTCATACAGGTGTCAGAAAGAGATGAAGTGAACATGATGGCTAAG 540
DB      481 GTTTCCTGTTACCTCATACAGGTGTCAGAAAGAGATGAAGTGAACATGATGGCTAAG 540
QY      541 CAAGAAAGCTAAGCTGTGCTCATGATGAAGAAATCAAGACCAAGCACTACATGATGAGATGTT 600
DB      541 CAAGAAAGCTAAGCTGTGCTCATGATGAAGAAATCAAGACCAAGCACTACATGATGAGATGTT 600
QY      601 ATTAGAATTTGTGCCAAGAAAGCAAGGCTCAGATCAATGCTTAAACGTTAACCA 660
DB      601 ATTAGAATTTGTGCCAAGAAAGCAAGGCTCAGATCAATGCTTAAACGTTAACCA 660
QY      661 GATGATCATGCGAGGAATTTCTCAAGAGTCTTGAAGAGAGATGATGATGACAAGTTC 720
DB      661 GATGATCATGCGAGGAATTTCTCAAGAGTCTTGAAGAGAGATGATGATGACAAGTTC 720
QY      721 CTGGAACCTTTGAGGTCAACCATCAGTCTTGAACAAGACAGAGCTTAACTTGTGCAT 780
DB      721 CTGGAACCTTTGAGGTCAACCATCAGTCTTGAACAAGACAGAGCTTAACTTGTGCAT 780
QY      781 GTTCTTCGTTGACGATCAACAAACTGGAATCTGATGAAGGAGCACTCACTGAATTGTG 840
DB      781 GTTCTTCGTTGACGATCAACAAACTGGAATCTGATGAAGGAGCACTCACTGAATTGTG 840
QY      841 ACCACAAGAGCTGAGATTGACTTGAAGGCTTGAAGAGAGATGATGATGACAAGAGCTC 900
DB      841 ACCACAAGAGCTGAGATTGACTTGAAGGCTTGAAGAGAGATGATGATGACAAGAGCTC 900
QY      901 ATTCTTTGAGAAAGCTATTACCAAGACACTCTGTGAGATTACAGAAAGATGCTGCTC 960
DB      901 ATTCTTTGAGAAAGCTATTACCAAGACACTCTGTGAGATTACAGAAAGATGCTGCTC 960
QY      961 GCACCTTCGCGTGAAGATGATGCTTAATCAATCAATCTTCCACAGAGAAACATAAGCTGC 1020
DB      961 GCACCTTCGCGTGAAGATGATGCTTAATCAATCAATCTTCCACAGAGAAACATAAGCTGC 1020

```

```

DB      961 GCACCTTCGCGTGAAGATGATGCTTAATCAATCAATCTTCCACAGAGAAACATAAGCTGC 1020
QY      1021 TCTACAGCTTCTGTATCTTATCTCCCTCTCTCTTGTGATGAGTTCAATTCGT 1080
DB      1021 TCTACAGCTTCTGTATCTTATCTCCCTCTCTCTTGTGATGAGTTCAATTCGT 1080
QY      1081 TGATTTGTTTCTACAAAACCTTGTGTTTCTGTGTGTGTTTGAATTCCTAAATTA 1140
DB      1081 TGATTTGTTTCTACAAAACCTTGTGTTTCTGTGTGTGTTTGAATTCCTAAATTA 1140
QY      1141 TGCAAAAGAGAGACAGAGAGAACCAAGTGTGCTCTTAAGTTATATATATGAAGAG 1200
DB      1141 TGCAAAAGAGAGACAGAGAGAACCAAGTGTGCTCTTAAGTTATATATATGAAGAG 1200
QY      1201 CATGGCCCTTAAAAAATTTTAAAAA 1230
DB      1201 CATGGCCCTTAAAAAATTTTAAAAA 1230

```

RESULT 2  
 AY072347  
 LOCUS  
 DEFINITION  
 Arabidopsis thaliana Ca2+-dependent membrane-binding protein  
 annexin (Atlg35720) mRNA, complete cds.  
 ACCESSION  
 AY072347  
 VERSION  
 AY072347.1 GI:18252242  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Arabidopsis thaliana (chale crese)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis  
 1 (bases 1 to 1208)  
 AUTHORS  
 Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,  
 Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,  
 Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,  
 Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,  
 Sakano,H., Sakurai,T., Saitou,M., Seki,M., Shim,P., Yamada,K.,  
 Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.  
 DIRECT SUBMISSION  
 Submitted (02-JAN-2002) DNA Sequencing and Technology Center,  
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
 USA  
 COMMENT  
 e-mail for correspondence: arab@sequence.stanford.edu

TITLE  
 JOURNAL  
 RIKEN Genomic Sciences Center (GSC) members carried out the  
 collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
 Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,  
 Saitou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
 Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, RGC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAFL cDNAs: Nguyen,M.,  
 Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,  
 Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,  
 Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shim,P.,  
 Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed  
 equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.  
 (SSP/Stanford) contributed equally to this work as PIs.

location/Qualifiers

## FEATURES

1..1208

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="1"

/clone="RAFL06-11-P04"

/ecotype="Columbia"

/note="This clone is in pBluescript"

1..1208

gene

33..986

CDS

```
/gene="AL1935720"
/codon_start=1
/product="Ca2+-dependent membrane-binding protein annexin"
/db_xref="GI:18252243"
/translation="MATLKVSDVPAPSDAEQRTAFEGMGTEEDLLISILAHREAE
ORVYIAQHEHYEDDLTKLDEKLSNDERAILMTWLEGEEDALLANEARMTSS
NOVIMEVACTRTSTOLLHARAYHARVYKSLIEDVANHHTGGPRKLLVSLNYSYREG
DEVNMTLAKQEKLVHEKIKDKRYNDEEDVIRIISTRSKQIINATFRYDDHSEELK
SLEEGDDKFLALRSTIQCLTRPELYVDVLRSAINTGTGALTRI VYTRAEID
LKVIGREYORRNSIPLERKAITKDTREGDYERKMLVALLGEEDA"
```

## ORIGIN

```
Query Match      97.3%; Score 1197.4; DB 8; Length 1208;
Best Local Similarity 99.9%; Pred. No. 1,5e-292;
Matches 1198; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 10 CGAACAACATAAAGTGAAGAAAAATGCGACCTTAAGGTTCTGATTTCTGCTGCT 69  
| | | | |  
DB 9 CAAAACCTRAAAGTGAAGAAAAATGCGACCTTAAGGTTCTGATTTCTGCTGCT 68  
| | | | |  
QY 70 CCTTGTGATGATGCTGAGCAATTGAAACCGCTTTTGAAGATGGGCTACGAAGAC 129  
| | | | |  
DB 69 CCTTGTGATGATGCTGAGCAATTGAAACCGCTTTTGAAGATGGGCTACGAAGAC 128  
| | | | |  
QY 130 TTGATCATATCATCTTTGGCTCACAGAGTGTGAACAGAGAAAGTCATCAGGCAAGCA 189  
| | | | |  
DB 129 TTGATCATATCATCTTTGGCTCACAGAGTGTGAACAGAGAAAGTCATCAGGCAAGCA 188  
| | | | |  
QY 190 TACCAAGAAACCTACGCGGCAAGACCTTCTCAAGACTTTGACAAGAGCTCTCAAGCAT 249  
| | | | |  
DB 189 TACCAAGAAACCTACGCGGCAAGACCTTCTCAAGACTTTGACAAGAGCTCTCAAGCAT 248  
| | | | |  
QY 250 TTGAGAGAGCTATCTTGTGTGGAATCTTGAACCCGGTGAAGGCTGATGCTTTATTTGGCT 309  
| | | | |  
DB 249 TTGAGAGAGCTATCTTGTGTGGAATCTTGAACCCGGTGAAGGCTGATGCTTTATTTGGCT 308  
| | | | |  
QY 310 AATGAAGCTACAAAAAGATGAGATTTCAAGACCAAGATTCTTATGGAAGTGTCTTGCA 369  
| | | | |  
DB 309 AATGAAGCTACAAAAAGATGAGATTTCAAGACCAAGATTCTTATGGAAGTGTCTTGCA 368  
| | | | |  
QY 370 AGGACATCAACGAGCTGCTTACCGGCAAGCTTACCATGCTGCTACAGAAAGTCT 429  
| | | | |  
DB 369 AGGACATCAACGAGCTGCTTACCGGCAAGCTTACCATGCTGCTACAGAAAGTCT 428  
| | | | |  
QY 430 CTTGAAGAGAGCGTTGCTCACACACTACCGGCTGACTTCAAGAAAGCTTTTGTTCCTT 489  
| | | | |  
DB 429 CTTGAAGAGAGCGTTGCTCACACACTACCGGCTGACTTCAAGAAAGCTTTTGTTCCTT 488  
| | | | |  
QY 490 GTTACTCTATCAAGGTCAGAGAGATGAAATGAATGAACTGCTGCTACAGAAAGCT 549  
| | | | |  
DB 489 GTTACTCTATCAAGGTCAGAGAGATGAAATGAATGAACTGCTGCTACAGAAAGCT 548  
| | | | |  
QY 550 AAGCTGTTCATGAGAAATCAAGCAAGCACTACAGATGATGAGATGTTATTAGAAATC 609  
| | | | |  
DB 549 AAGCTGTTCATGAGAAATCAAGCAAGCACTACAGATGATGAGATGTTATTAGAAATC 608  
| | | | |  
QY 610 TTGTCCACAAGAGCAAGGCTCAGATCAATGCTATCTTTTAAACCGTTACCAAGATGATCAT 669  
| | | | |  
DB 609 TTGTCCACAAGAGCAAGGCTCAGATCAATGCTATCTTTTAAACCGTTACCAAGATGATCAT 668  
| | | | |  
QY 670 GGGGAGAAATTTCAAGAGCTTTGAGAGAGAGATGATGATCAAGATTCCTTGAACCTT 729  
| | | | |  
DB 669 GGGGAGAAATTTCAAGAGCTTTGAGAGAGAGATGATGATCAAGATTCCTTGAACCTT 728  
| | | | |  
QY 730 TTGAGGTCAACCATTAAGGCTTGAACAAGCAAGAGCTTTACTTGTGCAATGTTCTTGGT 789  
| | | | |  
DB 729 TTGAGGTCAACCATTAAGGCTTGAACAAGCAAGAGCTTTACTTGTGCAATGTTCTTGGT 788  
| | | | |  
QY 790 TCAGCAATCAACAAAACCTGGAATGATGAGAGCACTCACTAAGATTTGTGACCAAGA 849  
| | | | |  
DB 789 TCAGCAATCAACAAAACCTGGAATGATGAGAGCACTCACTAAGATTTGTGACCAAGA 848  
| | | | |

QY 850 GCGGAGATTGACTTGAAGGTCATTTGAGAGAGATCAAGCGGAGAACGATTCCTTG 909  
| | | | |  
DB 849 GCTGACATTGACTTGAAGGTCATTTGAGAGAGATCAAGCGGAGAACGATTCCTTG 908  
| | | | |  
QY 910 GAGAAAGCTATTATCCAAAGACACCTGCGAGATTTACGAGAAAGTCTGCTGCATCTTC 969  
| | | | |  
DB 909 GAGAAAGCTATTATCCAAAGACACCTGCGAGATTTACGAGAAAGTCTGCTGCATCTTC 968  
| | | | |  
QY 970 GGTGAAGATGATGCTTAAATCAATCAATCTCCACAGAGAAACATAGTCTCTACAGCT 1029  
| | | | |  
DB 969 GGTGAAGATGATGCTTAAATCAATCAATCTCCACAGAGAAACATAGTCTCTACAGCT 1028  
| | | | |  
QY 1030 TCTGTTATCTTATCTCCCTCTCTCTCTCTCTTGTGATGAGTTCAATGCTTTGATTTGT 1089  
| | | | |  
DB 1029 TCTGTTATCTTATCTCCCTCTCTCTCTCTCTTGTGATGAGTTCAATGCTTTGATTTGT 1088  
| | | | |  
QY 1090 TTCTACAAAACCTTGTGTTGTTCTGTTGTTGTTGTTGATTCTTAATATGCAAAAG 1149  
| | | | |  
DB 1089 TTCTACAAAACCTTGTGTTGTTCTGTTGTTGTTGATTCTTAATATGCAAAAG 1148  
| | | | |  
QY 1150 GAGAGACAGAGAAACCAAGTGTGCTCTTAAGTATATATATATGAGAGCAATTGGCC 1208  
| | | | |  
DB 1149 GAGAGACAGAGAAACCAAGTGTGCTCTTAAGTATATATATATGAGAGCAATTGGCC 1207  
| | | | |

RESULT 3  
AY086570 1192 bp mRNA linear PLN 14-APR-2003  
LOCUS AY086570  
DEFINITION Arabidopsis thaliana clone 25846 mRNA, complete sequence.  
ACCESSION AY086570  
VERSION AY086570.1 GI:21405280  
KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana

REFERENCE 1 Haas, B. J., Volkovskiy, N., Town, C. D., Troukhan, M., Alexandrov, N.,  
Bukaryova, K. A., Flavell, R. B., White, O. and Salzberg, S. L.  
Full-length messenger RNA sequences greatly improve genome  
annotation.

JOURNAL Genome Biol. 3 (6), RESEARCH0029 (2002)  
MEDLINE 22088475  
PUBMED 12093376

REFERENCE 2 (bases 1 to 1192)  
AUTHORS Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and  
Feldmann, K.

TITLE Full-length cDNA from Arabidopsis thaliana  
JOURNAL Unpublished  
AUTHORS 3 (bases 1 to 1192)

TITLE Direct SubMISSION  
JOURNAL Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,  
Malibu, CA 90265, USA

COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made  
available to TIGR and Genbank. The following quality assessment of  
this set was done by comparison with known proteins: two percent  
of the clones are estimated to be 5'-truncated; less than one percent  
are 3'-truncated; approximately two percent represent alternative  
splice variants, including unspliced introns and spliced exons; one  
percent may contain premature stop codons; five percent may have  
frame shifts in a coding region. A sequence is considered to be  
5'-truncated if it lacks the translation initiation start (ATG). A  
sequence is considered to be 3'-truncated if it lacks the  
C-terminal end of the encoded protein. Please note that these cDNA  
sequences are derived from the Ws or Laer ecotypes and therefore  
may contain polymorphisms when compared to sequences from Col-0.  
Genest carried out the library production and sequences from Col-0.  
full-length clones. Ceres, Inc. carried out the clustering of the  
5' sequences, selection of clones, and sequence assembly.

## FEATURES

Location/Qualifiers

source	1. .1192
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/db_xref="taxon:3702"
	/clone="25846"
CDS	36. .989
	/codon_start=1
	/product="Ca2+-dependent membrane-binding protein annexin"
	/protein_id="AA063633.1"
	/db_xref="GI:21554612"
	/translation="MATTAKVSDSPAPSDAQLFTAPRGKTNEDLIISILAHSAE QKVIHQAYHETVGDLEKTLKESLNDERAIIIMWLEPERDALLNVAEKRTSS NQVMEVACTRTSTQLHARQAYHARVKKSLSDVAHHTTGDPRKILVSLVSYVEG DEVMNTLAKQEAHLVHEKIKDHYNDEVDVIRLSTRSAQINATVGVODDSEILK SLREGDDDKFLALRSTIGCTREPLVDVLRSAINTGTDEGALRIVYTRAID LKVIEEYORRNSIPLEKAITDTRGDVHKMLVALLGEDA"
ORIGIN	
Query Match	94.5%; Score 1162.4; DB 8; Length 1192;
Best Local Similarity	99.1%; Pred. No. 1.1e-283;
Matches 1169; Conservative	0; Mismatches 11; Indels 0; Gaps 0;
QY	10 CGAAACACTAAAGTGAAGAAAATGGGACTCTTAAGCTTTCTGATTCCTGCT
DB	12 CAAACACTAAAGTGAAGAAAATGGGACTCTTAAGCTTTCTGATTCCTGCT
QY	70 CCTTCGATGATGCTGAGCAATTTGAGAACCGCTTTTGAAGGATGGGGTACGACGAGC
DB	72 CCTTCGATGATGCTGAGCAATTTGAGAACCGCTTTTGAAGGATGGGGTACGACGAGC
QY	130 TTGATCATATCAATCTTGGCTCAGAGAGTGTGAACAGAGGAAGTCATCAGGCAAGCA
DB	132 TTGATCATATCAATCTTGGCTCAGAGAGTGTGAACAGAGGAAGTCAGGCAAGCA
QY	190 TACCAAGAACTTACGGCGAAGACCTTTCAGAGACTCTTGACAGAGAGCTCTTAACAT
DB	192 TACCAAGAACTTACGGCGAAGACCTTTCAGAGACTCTTGACAGAGAGCTCTTAACAT
QY	250 TTGAGAGAGCTATCTTGTGTGAGACTCTTGAACCGGAGCGTGAATCTTATATGGCT
DB	252 TTGAGAGAGCTATCTTGTGTGAGACTCTTGAACCGGAGCGTGAATCTTATATGGCT
QY	310 AATGAAGCTACAAAAGATGACTTCAAGCAACCAAGTCTTATGAAAGTTGCTTGACA
DB	312 AATGAAGCTACAAAAGATGACTTCAAGCAACCAAGTCTTATGAAAGTTGCTTGACA
QY	370 AGGACATCAACGACGCTGTTACGCTGAGGCAAGCTTACATGCTGCTCAACAAGTCT
DB	372 AGGACATCAACGACGCTGTTACGCTGAGGCAAGCTTACATGCTGCTCAACAAGTCT
QY	430 CTGGAAGAGAGCTTGGCTCAGCACTACCGGTGAAGCTTCAAGAAAGCTTTGGTTCTCT
DB	432 CTGGAAGAGAGCTTGGCTCAGCACTACCGGTGAAGCTTCAAGAAAGCTTTGGTTCTCT
QY	490 GTTACCTCATACAGTACGAGAGAGATGAAGTGAACATGACATTTGGCTTAAGCAAGAGCT
DB	492 GTTACCTCATACAGTACGAGAGAGATGAAGTGAACATGACATTTGGCTTAAGCAAGAGCT
QY	550 AAGCTGTTCATGAGAAAATCAAGACAGCACTACATGATGAGAGATTTTGAATC
DB	552 AAGCTGTTCATGAGAAAATCAAGACAGCACTACATGATGAGAGATTTTGAATC
QY	610 TTGTCACAAAGAGCAAGCTCAGATCAATGTCTACTTTAAACGGTTTCCAAATGATGAT
DB	612 TTGTCACAAAGAGCAAGCTCAGATCAATGTCTACTTTTAAACGGTTTCCAAATGATGAT
QY	670 GCGAGGAAATTTCTCAAGAGCTTTGAAGAGAGATGATGATGACAAAGTTCTTGCACAT
DB	672 GCGAGGAAATTTCTCAAGAGCTTTGAAGAGAGATGATGATGACAAAGTTCTTGCACAT
QY	730 TTGAGCTCAACCATTCAGTCTTGACAGACAGAGCTTACTTTGTGATGATTTCTTGT
DB	732 TTGAGCTCAACCATTCAGTCTTGACAGACAGAGCTTACTTTGTGATGATTTCTTGT

QY	790 TCAGCAATCAACAAAATGGAACCTGATGAGAGACACTCACTAGATTTGTGACACAAGA	849
DB	792 TCAGCAATCAACAAAATGGAACCTGATGAGAGAGACACTCACTAGATTTGTGACACAAGA	851
QY	850 GCTGAGATTGACTTGAAGGTCATTGGAGAGAGATGACAGCCGAGGACAGCATTCCTTTG	909
DB	852 GCTGAGATTGACTTGAAGGTCATTGGAGAGAGATGACAGCCGAGGACAGCATTCCTTTG	911
QY	910 GAGAAAGCTATTACCAAGACACTCGGTGAGATTCGAGAAAGATGCTGTCACATTC	969
DB	912 GAGAAAGCTATTACCAAGACACTCGGTGAGATTCGAGAAAGATGCTGTCACATTC	971
QY	970 GGTGAAGATGATGCTTAATCAATCAATCTCCACAGAGAAACATTAAGCTGCTACAGCT	1029
DB	972 GGTGAAGATGATGCTTAATCAATCAATCTCCACAGAGAAACATTAAGCTGCTACAGCT	1031
QY	1030 TCTGTTATCTCTTAATCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1089
DB	1032 TCTGTTATCTCTTAATCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1091
QY	1090 TTCTACAAAACCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1149
DB	1092 TTCTACAAAACCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1151
QY	1150 GAGAGACAGAGAGAACAGCAGTGTGCTCTTAAGTTATATA	1189
DB	1152 GAGAGACAGAGAGAACAGCAGTGTGCTCTTAAGTTATATA	1191

RESULT 4	BT003359	1139 bp	mRNA	linear	PLN 29-JAN-2003
LOCUS	BT003359				
DEFINITION	Arabidopsis thaliana Ca2+-dependent membrane-binding protein				
ACCESSION	BT003359				
VERSION	BT003359.1	GI:28059005			
KEYWORDS	FLI CDNA.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Southwick, A., Nguyen, M., Tripp, M., Palm, C.J., Jones, T., Wu, T., Deng, J.M., Hayashizaki, Y., Hsuan, V.W., Lee, J.M., Ishida, J., Kamliya, A., Kawai, J., Kim, C.J., Narusaka, M., Onodera, C.S., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Shimizu, P., Tang, C.C., Torouni, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.				
AUTHORS	Direct Submission				
TITLE	Submitted (29-JAN-2003) DNA Sequencing and Technology Center,				
JOURNAL	Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA				
COMMENT	e-mail for correspondence: arab@sequence.stanford.edu				

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : "RIKEN Arabidopsis Full-Length CDNA") : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamliya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen, M., Tripp, M., Southwick, A., Palm, C.J., Jones, T., Wu, T., Chen, H., Cheuk, R., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Kim, C.J., Quach, H.L., Onodera, C.S., Shimizu, P., Tang, C.C., Torouni, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Ecker, J., Theologis, A. and Davis, R.W.

Nguyen, M., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W.





/db\_xref="UniProt/TREMBL:Q39001"  
 /translation="TLKSDVSPAPSDDAEOLRTAFEGMGTNEDLIISILAHSAEOR  
 KYRQAYHETGSDILKTDKLSNDFEALILMTLEPERAILANETKRWTSNO  
 VMEVACTRTSTDLHARQAYHARKYKSLBEDVAHHTTSDFKLVLSTVSTRYBDE  
 VNMVTLAKQEKLVHEKIKDKHYNDEVIVILSTISKAQINATFNRRQDDHGEILKSL  
 EEDDDDKPECTLRSTIOCTRELYFVNLRSAINKTGTDEGALTRIVTTAEIDLK  
 VIGEEYQRNSIPLRAITKDNCGDYKMLVAFGEEDA"

## ORIGIN

Query Match 91.4%; Score 1124.2; DB 8; Length 1206;  
 Best Local Similarity 97.0%; Pred. No. 5,5e-274;  
 Matches 1157; Conservative 0; Mismatches 33; Indels 3; Gaps 1;

38 CGACTTTAAGGTTTGTGATTTCTGCTCGTCTCTTCTGATGAGTGAAGATTTAGAA 97  
 1 CGACTTTAAGGTTTGTGATTTCTGCTCGTCTCTTCTGATGAGTGAAGATTTAGAA 60  
 98 CCGCTTTTGAAGATGGGGTACGACGAGGACTTGATCATATCATCTTGGCTCAGAGAA 157  
 61 CCGCTTTTGAAGATGGGGTACGACGAGGACTTGATCATATCATCTTGGCTCAGAGAA 120  
 158 GTGCTGAACGAGAGAAAGTATCAGGCAAGCATACCAAGAACTTACGCGAAGACTTTC 217  
 121 GTGCTGAACGAGAGAAAGTATCAGGCAAGCATACCAAGAACTTACGCGAAGACTTTC 180  
 218 TCAAGACTCTTGAAGAAGCTCTTACAGATTTGAGAGAGCTATCTTGTGTGAGACTC 277  
 181 TCAGAGCTCTTGAAGAAGCTCTTACAGATTTGAGAGAGCTATCTTGTGTGAGACTC 240  
 278 TTGAACCCGGTGAAGCGTATGCTTTTATGGCTAATAGAAAGCTACAAAAGATGACTTCA 337  
 241 TTGAACCCGGTGAAGCGTATGCTTTTATGGCTAATAGAAAGCTACAAAAGATGACTTCA 300  
 338 GCAACCAAGTTCTTATGGAAGTTGCTTGCACAAGGACATTAAGCGAGCTTCAAGCTTA 397  
 301 GCAACCAAGTTCTTATGGAAGTTGCTTGCACAAGGACATTAAGCGAGCTTCAAGCTTA 360  
 398 GCGAAGCTTACAGATGCTGCTTACAGAAGATCTCTTGAAGAGAGCTTGTCTCACACACTA 457  
 361 GCGAAGCTTACAGATGCTGCTTACAGAAGATCTCTTGAAGAGAGCTTGTCTCACACACTA 420  
 458 CCGGTGACTTCAAGAAAGCTTTTGGTTTCTTGTCTTACCTTATACAGATTAAGAGAGATG 517  
 421 CCGGTGACTTCAAGAAAGCTTTTGGTTTCTTGTCTTACCTTATACAGATTAAGAGAGATG 480  
 518 AATTGAACATGATGATTTGCTTAAGCAAGAGCTTAAGCTGCTCCATGAGAAATTAAGAGCA 577  
 481 AAGTGAACATGATGATTTGCTTAAGCAAGAGCTTAAGCTGCTCCATGAGAAATTAAGAGCA 540  
 578 AGCACTACATGATGATGATTTATTAAGATCTTGTCCACAAGAACCAAAAGCTCAGATCA 637  
 541 AGCACTACATGATGATGATTTATTAAGATCTTGTCCACAAGAACCAAAAGCTCAGATCA 600  
 638 ATGCTACTTTTAAACCGTTACCAAGATGATCATGCGCGAGAAATTTCCAGAAGCTTGAAG 697  
 601 ATGCTACTTTTAAACCGTTACCAAGATGATCATGCGCGAGAAATTTCCAGAAGCTTGAAG 660  
 698 AAGAAGATGATGATGATTTCTTGCACATTTTGAAGTAAACCATTTCACTGCTTTGACAA 757  
 661 AAGAAGATGATGATGATTTCTTGCACATTTTGAAGTAAACCATTTCACTGCTTTGACAA 720  
 758 GACCAAGAGCTTACTTGTGATGATTTCTTGCACATTTTGAAGTAAACCATTTCACTGCTTTG 817  
 721 GACCAAGAGCTTACTTGTGATGATTTCTTGCACATTTTGAAGTAAACCATTTCACTGCTTTG 780  
 818 AAGAGAGCACTACTAGATTGTGACCAAGAGCTGAGATTGACTTGAAGCTATTGGAG 877  
 781 AAGAGAGCACTACTAGATTGTGACCAAGAGCTGAGATTGACTTGAAGCTATTGGAG 840  
 878 AGAGATACCAAGCGCAGAGAACAGATTTCTTGAAGAAAGTATTTACCAAGAACACTCGTG 937  
 841 AGAGATACCAAGCGCAGAGAACAGATTTCTTGAAGAAAGTATTTACCAAGAACACTCGTG 897

QY 938 GAGATTACGAGAGATGCTGCTGCGACTTCTGCTGAGAGATGATCTTAATCAATCAATC 997  
 DB 898 GAGATTACGAGAGATGCTGCTGCGACTTCTGCTGAGAGATGATCTTAATCAATCAATC 957  
 QY 998 CTCACAGAGAAACATTAAGCTGCTCTTACAGCTTCTGTTATCTCTTATCTCCCTCTCTC 1057  
 DB 958 CTCACAGAGAAACATTAAGCTGCTCTTACAGCTTCTGTTATCTCTTATCTCCCTCTCTC 1017  
 QY 1058 TCTTTGATGAGTTCAAAATCGTTGATTTTGTCTTACAAAACCTGTTGTTCTGT 1117  
 DB 1018 TCTTTGATGAGTTCAAAATCGTTGATTTTGTCTTACAAAACCTGTTGTTCTGT 1077  
 QY 1118 GTGTGTTTGAAGTCTTAATATATGCAAAAGAGAGACAGAGAACAGTGTGCTC 1177  
 DB 1078 GTGTGTTTGAAGTCTTAATATATGCAAAAGAGAGACAGAGAACAGTGTGCTT 1137  
 QY 1178 TTAAGTTATATATATGAGAGCACTTGGCTTAATATATATATATATATATATATAT 1230  
 DB 1138 TTAAGTTAT 1190

RESULT 6  
 ATU28415 1159 bp mRNA linear PLN 21-OCT-1996  
 LOCUS Arabidopsis thaliana annexin-like protein mRNA, complete cds.  
 DEFINITION U28415  
 ACCESSION U28415  
 VERSION U28415.1 GI:1621538  
 KEYWORDS  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
 1 (bases 1 to 1159)  
 Gidrol,X., Sabelli,P.A., Fern,Y.S. and Kush,A.K.  
 Annexin-like protein from Arabidopsis thaliana rescues delta oxyR  
 mutant of Escherichia coli from H2O2 stress  
 Proc. Natl. Acad. Sci. U.S.A. 93 (20), 11268-11273 (1996)  
 MEDLINE 97008169  
 PUBMED 8853345  
 2 (bases 1 to 1159)  
 Kush,A.K.  
 Direct Submission  
 Submitted (06-JUN-1995) Anil K. Kush, Institute of Molecular & Cell  
 Biology, National University of Singapore, 10 Kent Ridge Crescent,  
 Singapore 0511, Republic of Singapore

FEATURES  
 source location/Qualifiers  
 1..1159  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /variety="ecotype Columbia"  
 /db\_xref="taxon:3702"  
 /clone="pOxy5"  
 /cisse\_type="whole seedling"  
 /clone\_lib="Lambda Unizapir cDNA library of A.K.Kush"  
 /dev\_stage="4 week-old"  
 6..959  
 /note="annexin-like protein"

CDS  
 /codon\_start=1  
 /protein\_id="AAC49472.1"  
 /db\_xref="GI:1621539"  
 /translation="MATLKVDSPAPSDDAEOLRTAFEGMGTNEDLIISILAHSAE  
 ORKVIROYHETGSDILKTDKLSNDFEALILMTLEPERAILANETKRWTSNO  
 NOVMEVACTRTSTDLHARQAYHARKYKSLBEDVAHHTTSDFKLVLSTVSTRYBDE  
 DEVNMTLAKQEKLVHEKIKDKHYNDEVIVILSTISKAQINATFNRRQDDHGEILK  
 SLEEDDDDKPECTLRSTIOCTRELYFVNLRSAINKTGTDEGALTRIVTTAEID  
 LKVIIEEYQRNSIPLRAITKDNCGDYKMLVAFGEEDA"

## ORIGIN

Query Match 91.1%; Score 1120; DB 8; Length 1159;  
 Best Local Similarity 98.9%; Pred. No. 6,4e-273;  
 Matches 1149; Conservative 0; Mismatches 10; Indels 3; Gaps 2;

QY 29 GAAAAATGGGACCTTAAAGTTCTGATTCGTCTGCTCTTCTGATGATGCTAGC 88  
 DB 1 GAAAAATGGGACCTTAAAGTTCTGATTCGTCTGCTCTTCTGATGATGCTAGC 60  
 QY 89 AATTGAGAACCGCTTTTGAAGATGGGGTACGAACGAGACTTGATCATATCAATCTTGG 148  
 DB 61 AATTGAGAACCGCTTTTGAAGATGGGGTACGAACGAGACTTGATCATATCAATCTTGG 120  
 QY 149 CTCACAGAGTGTCTGAACAGAGAAAGTCATCAGGCAACATACCGAAACCTACGCGC 208  
 DB 121 CTCACAGAGTGTCTGAACAGAGAAAGTCATCAGGCAACATACCGAAACCTACGCGCA 180  
 QY 209 AAGACTTCTCAGAGCTCTTGACAAAGGAGCTCTTAACGATTTGAGAGAGCTATCTTGT 268  
 DB 181 AAGACTTCTCAGAGCTCTTGACAAAGGAGCTCTTAACGATTTGAGAGAGCTATCTTGT 240  
 QY 269 TGTGACTCTTGAACCCGGTGGAGGTGATGCTTTATTTGGCTTAATGAAGCTACAAAAGAT 328  
 DB 241 TGTGACTCTTGAACCCGGTGGAGGTGATGCTTTATTTGGCTTAATGAAGCTACAAAAGAT 300  
 QY 329 GGACTTCAAGCAACCAAGTTCTTATGGAAGTTCCTTGACAAGGACATCAACGACGTGC 388  
 DB 301 GGACTTCAAGCAACCAAGTTCTTATGGAAGTTCCTTGACAAGGACATCAACGACGTGC 360  
 QY 389 TTCAAGCTAAGGCAAGCTTAACATGCTGCTACAGAAAGTCTCTGAAGAGAGAGCTTGTCTC 448  
 DB 361 TTCAAGCTAAGGCAAGCTTAACATGCTGCTACAGAAAGTCTCTGAAGAGAGAGCTTGTCTC 420  
 QY 449 ACCACACTAACCGGTGACTTCAGAAAGCTTTGGTTTCTCTGTTACCTCATACAGATGACG 508  
 DB 421 ACCACACTAACCGGTGACTTCAGAAAGCTTTGGTTTCTCTGTTACCTCATACAGATGACG 480  
 QY 509 AAGGAGATGAAGTGAACATGACATTTGGCTAAGCAAGAGCTAAGCTGTCCATGAGAAA 568  
 DB 481 AAGGAGATGAAGTGAACATGACATTTGGCTAAGCAAGAGCTAAGCTGTCCATGAGAAA 540  
 QY 569 TCAAGGACAAGCACTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628  
 DB 541 TCAAGGACAAGCACTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
 QY 629 CTCGATCAATGCTACTTTTAAACGTTACCAAGATGATGATGATGATGATGATGATGATGATG 688  
 DB 601 CTCGATCAATGCTACTTTTAAACGTTACCAAGATGATGATGATGATGATGATGATGATGATG 660  
 QY 689 GTTCTGAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 748  
 DB 661 GTTCTGAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
 QY 749 GCTTGACAAGACGAGCTTTAATTTGTCATGTTTCTGTCAGCAATCAACAAACCTG 808  
 DB 721 GCTTGACAAGACGAGCTTTAATTTGTCATGTTTCTGTCAGCAATCAACAAACCTG 780  
 QY 809 GAACTGATGAAGAGCACTCACTGATTTGTCACCAAGAGCTGATGATGATGATGATGATGATG 868  
 DB 781 GAACTGATGAAGAGCACTCACTGATTTGTCACCAAGAGCTGATGATGATGATGATGATGATG 840  
 QY 869 TCATTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 928  
 DB 841 TCATTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
 QY 929 AACTCTGATGAGATTAAGAGAAAGTGTCTGTCGACATTTCTGCTGGAAGATGATGATGATG 988  
 DB 901 AACTCTGATGAGATTAAGAGAAAGTGTCTGTCGACATTTCTGCTGGAAGATGATGATGATG 960  
 QY 989 CAATCAATCTCTCCACAGAGAAACATAGCTGCTTACAGCTTCTGTTATCTTTATC-CT 1048  
 DB 961 CAATCAATCTCTCCACAGAGAAACATAGCTGCTTACAGCTTCTGTTATCTTTATC-CT 1019  
 QY 1049 CTCTCTCTCTTCTTGAAGATTTCAAAATGTTGATTTGTTTCTACAAAAACCTTGT 1108  
 DB 1020 CCTCTCTCTCTTCTTGAAGATTTCAAAATGTTGATTTGTTTCTACAAAAACCTTGT 1079  
 QY 1109 GTTCTCTCTCTTCTTGAAGTTCCTAAATTAATGCAAAAAGAGAGACAGAGAACAG 1168

DB 1080 GTTCTCTCTCTTCTTGAAGTTCCTAAATTAATGCAAA--AGAGAGACAGAGAACAG 1137  
 QY 1169 TGTGCTCTTAAAGTATATAT 1190  
 DB 1138 TGTGCTCTTAAAGTATATAT 1159  
 RESULT 7  
 P14D7  
 LOCUS  
 DEFINITION  
 Sequence of BAC P14D7 from Arabidopsis thaliana chromosome 1,  
 complete sequence.  
 ACCESSION  
 AC021198  
 VERSION  
 AC021198.2  
 KEYWORDS  
 HTG.  
 SOURCE  
 Arabidopsis thaliana (thale cress)  
 ORGANISM  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 81513)  
 REFERENCE  
 AUTHORS  
 Liu,S.X., Chan,A., Sakano,H., Yu,G., Lee,J.M., Lenz,C., Pham,P.,  
 Toriumi,M., Chin,C., Chlou,J., Choi,E., Chung,M., Gonzalez,A.,  
 Homag,B., Liu,A., Vayberg,M., Alcafi,H., Brooks,S., Buehler,E.,  
 Chao,Q., Conn,J., Conway,A.B., Hansen,N.F., Johnson-Hopson,C.,  
 Khan,S., Kim,C., Lam,B., Miranda,M., Nguyen,M., Palm,C.J.,  
 Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R., Federspiel,N.A.,  
 Theologis,A. and Yu,G.  
 The sequence of BAC P14D7 from Arabidopsis thaliana chromosome 1  
 Unpublished  
 2 (bases 1 to 81513)  
 REFERENCE  
 AUTHORS  
 Theologis,A.  
 Direct Submision  
 Submitted (15-JAN-2000) Plant Gene Expression Center, 800 Buchanan  
 Street, Albany, CA 94710, USA  
 3 (bases 1 to 81513)  
 REFERENCE  
 AUTHORS  
 Theologis,A.  
 Direct Submision  
 Submitted (09-FEB-2000) Plant Gene Expression Center, 800 Buchanan  
 Street, Albany, CA 94710, USA  
 4 (bases 1 to 81513)  
 REFERENCE  
 AUTHORS  
 Theologis,A.  
 Direct Submision  
 Submitted (28-JUN-2000) Plant Gene Expression Center, 800 Buchanan  
 St., Albany, CA 94710, USA  
 On Feb 11, 2000 this sequence version replaced gi:6693723.  
 COMMENT  
 The sequence is of BAC P14D7 from Arabidopsis thaliana chromosome  
 1. The sequence does not represent the sequence of the entire  
 insert of this clone. It is shorter by 6954 bp because we submit  
 only the unique sequence of the clone. However, in order to  
 facilitate the joining of overlapping clones in the future for  
 creation of larger contigs, we provide small overlaps (200 bp)  
 between overlapping subitted clones. The 5' end of this sequence  
 overlaps by 200 bp to the 3' end of the sequence of the clone  
 F1504.  
 FEATURES  
 source  
 Location/Qualifiers  
 1. 81513  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /cultivar="Columbia"  
 /db\_xref="taxon:3702"  
 /chromosome="1"  
 /clone="P14D7"  
 1553..4999  
 /gene="P14D7.1"  
 join(1553..4628,4713..4999)  
 /gene="P14D7.1"  
 /note="Contains similarity to receptor protein kinase-1like  
 protein from Arabidopsis thaliana gb|AL161513. It  
 contains a eukaryotic protein kinase domain PF|00069. EST  
 gb|AI997574 comes from this gene."  
 /codon\_start=1

/evidence=not experimental  
/protein\_id="AA79882.1"  
/db\_xref="GI:8778966"  
/translation="MGFAEKNLVDFRPLFPISTIIISCSISASATIAEANAALKMNSTP  
TNSSKLSWHDANNTSPSCISWYGVNSNRSQISBELNVTGIEGTQDPFPIISLS  
NLAAYLISNMLSGTIPQFGNLSKLIYEDLSNHLTGEISPSGLNKLTVLYLHON  
YLSVYPSIEGNMSEMTDLASQNLKTSIIPSSGLNKLMLVLYEANTLTVIPEL  
GNMSEMTDLASQNLKTSIIPSTLGNLXNMLVLYEANTLTVIPEIENMSEMTNLA  
LSQNLKTSIPSSGLNKLTVLISLPOVLTGSIIPKLGNTSEMTDIEISNNKLTGS  
PSLGNLKNLTVIYLYEANTLTVIPELGNMSEMTDQNNKLTGSISPSGLNKL  
TVLYLYLNTLTVIPELGNMSEMTNLDLSQNLKTSVDSFGNFTKESLSGLNKL  
SGAIPPGVANSHTLTVLIDTNNFTGFPEPYCKAKLQNLISDYNHLEGP1PKSLR  
CSKLIARFELGNFTGDIPEAFGFIYDPLFIIDPSHAKFGEISNNMSEPKLGALMS  
NNITGAIPTIEINNMTOVLELDTNNLFGELPEAGLNTLSRLNINOLSGRYPA  
GISFLNLSLSDLSNNPSEIIPOTPSFLKLDPMVLSRNKPDGSIIPKLTCTOL  
DUSHNODGEIPIQSLSLOSLDKLDSHNNLSGLITTPTEGMALTNVDSINNKLEGP  
LPDPTPRKATADALEENIGLCSNIPKQRLKPCRELKRRKGNLVMLVLIYGLV  
LISICANTFTYCIKRRKQNGNTDETCENMSIFVDSKPYODIESTNEFDPHL  
IGTGYSKYRANLQDTIIAIVKRLHDTIDEISKPVKOEPLNEVKALTEIHRNVK  
LFGFSGHRHTFLIYEMKESLNKLANDSEAKRLTWTKRINNVGVAAHALSYMH  
RTPRIYHRDSSGNLIDNDYAKISDPGTAKLTKTGSNMSVAGTGYVAPAPAYT  
MKYTEKDVYSGVILIEILITIGHRPDVLVSSSSGFEALSRISDERVLSRQNR  
EKLLKNVEMALLCLQANPESRPTMLSITTPS"  
5917. .7552  
/gene="F14D7.2"  
/join(5917. .5992,6589. .6734,6821. .7552)  
/gene="F14D7.2"  
/note="identical to annexin (AnnAcl) mRNA from Arabidopsis  
thaliana gb|Af083913. It contains an annexin domain  
pf|00191. Bstse gb|H76460. gb|Z18518. gb|Z26190.  
gb|N96455. gb|Z47714. gb|T41940. gb|T43657. gb|N95995,  
gb|R30014. gb|T22046. gb|H37398. gb|H77008. gb|R29768,  
gb|H3260. gb|Z17514. gb|W43175. gb|T76739. gb|AA712753.  
gb|H76134. gb|T42209. gb|H36536. gb|AI1998553. gb|Z33565.  
gb|AA597533. gb|AI100145 and gb|AI100054 come from this  
gene."  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="AA79882.1"  
/db\_xref="GI:8778967"  
/translation="MATLVSDSVAPSDADQRLTAPEGCTNEDLIISIIANSAB  
ORVIRQAVHETYGEDLKLTDKELSENDEKRLIMLTLPGEREDLALNEAKRWSS  
NOYLMVACTRSTQHLHARQAVHAKYKSLSEVDHHTGDFRKLVSIVTSYREG  
DEVNMTLAKOEAQLVHEKIKDKHYNVEDYIRISSTRSKAONATENRVOYDDGEELK  
SLERGGDDDKFLALASTIOCLTRPELVYDVRASINKTGDEGALTRIVYTRAED  
LKVTGEYQRNSIDPEKAITKDTREDYKEMVALLGEEDA"  
/complement(7937. .10730)  
/gene="F14D7.3"  
/complement(join(7937. .8041,8115. .8364,8443. .8606,  
8757. .8919,9397. .10730))  
/gene="F14D7.3"  
/note="Contains similarity to a putative protein CAB78009  
gi|7267527 from Arabidopsis thaliana BAC T32A17  
gb|Al161513. It contains Pumilio-family RNA binding  
domains pf|00806."  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="AA79883.1"  
/db\_xref="GI:8778968"  
/translation="MKRONKRRKGSVNSSISAGCKQCLRCYDINGEYKTKETNT  
THSLRTYHATTFSLISLQTSICYITRPLSLSPYCFGLTSLSPFVCLVYFM  
FGRFPGFGFNGDFNFGASDHRPLGSDGFLPSLDPNPLKNNKHEVLAIVYFM  
MGISCDMSIWTKEPQDPRVDSRFGARLHSGSGPONLGAOLHDEGRNRSVAYO  
NNNFHVSPSPGEMRLRGODSPNLNGFEMALAKNRDPLDQIHETPKRPPPLRGN  
DALISLMPFEGNNRVSOQLAMASRGFTPEEDSLSLSPHEVSGKASVLEUETD  
IGGSGYKMSPKSNNDLVSWEIYGSVNLMAQDQICRYLQKVEGTGHEAKVILA  
IIDHVELSMDPGNYIVOKLIVGSDSEORLILVSLTNPPELLRICANTGYRVAQ  
KMEYVTKQOILVKSGLKPGFLAVLQNGHVGOSGLQTIQSPDNFVEAATKY  
CAEILHRRGCCVLOCCISNSVGLQREPLVAISIRSLSDSPENYVOVLLDOOV  
SAVKLVQFRMHAEALQKFSFHVIEKLCIKTPRESAIVRELLCVPEFVILLDDPY  
ANVITQIALSVTKGPVRAKLVAKVYVYGLHSSPYCKKILPSKILKK"  
12116. .14526  
/note="Contains similarity to Tal-1\_transposon gb|X13291"

gene	complement(19422. .23686) /gene="F14D7.4" /complement(join(19422. .19496,19746. .19835,20053. .20194, 23655. .23686)) /gene="F14D7.4" /note="Contains similarity to 8-amino-7-oxonanoate synthase-like protein gi 7406466 from Arabidopsis thaliana BAC T32A21 gb Al162875. Est gb AI994915 comes from this gene." /codon_start=1 /evidence=not experimental /protein_id="AA79884.1" /db_xref="GI:8778969" /translation="MAKKATVAIICFTLIGEIFSGVPAEGRKGRFKLLFESVNDL GLSSPTSLNAAALFCEGLIGCSSTSRKLSLKKHAKFTLEKVDADDEKHKNED EYGEDDDYKR" complement(34011. .36183) /gene="F14D7.5" complement(join(34011. .34115,34201. .34450,34532. .34788, 34873. .35035,35279. .36183)) /gene="F14D7.5" /note="Contains similarity to a putative RNA-binding protein AA618550 gi 6587864 from Arabidopsis thaliana BAC T1111 gb AC012680. It contains Pumilio-family RNA binding domains pf 00806." /codon_start=1 /evidence=not experimental /protein_id="AA79885.1" /db_xref="GI:8778970" /translation="MEIFNFGASDRRLPDFSGGFIQSLDITNPFLNKQYNNVSEA LELCKRLKNNKISCDMSIWTKEPDEFPDPPGAKTLHESRFGDONLTGASQIHDF RNFSSVRQNNNFHGVSPSGLLGLQDSFNGEEMMAFDQHDINERPKRS PLRNDAPKGSIMPEGIRVSOILAAHSGASVDEPKINGGLDLSVAEIVYSV NLMARDQGCALQKLVGEGYLDKSLVLEIIDHYVLSLDPGNTYVOVLVAVSDE EORTNIVLTSKREBELKICLANTGTRVITOMIKTYTKQIALVKSALBPGFLV NDSNGYHVLQCLFELVNDKRLRTKLKSHLISFLPLTVSSDCECEICQFVEA ATEYCAOLATHQGCYVLCSSLINTVGLQHRVLVAEISRDSLRUSDPPGNYVQCL DOOVSVNVLPLFRTHCIELATOKFSFHVIEKLCIKTPRESAIVRELSYDNFQQL DDPYANVITQIALSVTKAVARALVEKTKRGRKGLQSNYCKKIFSKILKK" 44612. .48698 /gene="F14D7.6" /join(44612. .44780,45014. .45403,45501. .45644,45676. .46134, 46297. .46476,46633. .47029,47113. .47505,47737. .48082, 48145. .48411,48465. .48641,48632. .48698) /gene="F14D7.6" /note="Strong similarity to a mutator-like transposase gi 4063759 from Arabidopsis thaliana BAC T14A4 gb AC005561." /codon_start=1 /evidence=not experimental /protein_id="AA79886.1" /db_xref="GI:8778971" /translation="MSDLYGVARHLYTVLLAVYSLCLKEKPRKLTVSDKINA KGAKRPNRSRLKVARLRQDGFPAKVAVKELLAGADDIELTYOWQMMGPDWR ANPILIDDEDETFMAIRADLDEHVPVKIRGNSNNVSYSNLDIGLTLTEIS DKYNSAETRAVWGTLRMLRLNRNSAIIISPIYLYATVILYIQVNDQMPWERTTA MAYRMANNTMETMGAGTSGVPLTDPSPDSIPISKGEKYLDDERBRDLPEMWT QOQEBKREKCATQWTLGATQIDATNVSDTETKARALNFGSGDIFKRNPRDLDTLA RAETVECTQVVALSSSDSTFGMPISLTSVSFSDDLMDVTDSSNNSSCTIDGKG HMLIDRITTKKEIAGASTNTRVSEYLRNLEMLDPPVPLVLDROAPVFDPEKSDAL HMLVYANBEGDIIYIGRIFKSKEBCAKFLAIIHIMHFHRIYASGNPIVAACVSD TCPMPVYATKLPDSMPFVRSSTYQNGCSVARGDFHQAQSTVVGKLMRTYIAGVR GPRANLEBRLMRODVALIOPYFKCLLETNPNSIPLVMEKQNSGVREKRYLFEALDAC VQGYAMKRVIVIDGTHLRKRYGGGLVAASQDANFYQPLPAFGLVNSKNDKEMATWFM
gene	Query Match Best Local Similarity 92.2%; Pred. No. 1.1e-246; Matches 1117; Conservative 8; Indels 86; Gaps 1;
CDS	102 TTTTGAAGATGGGTAAGAACAGAGACTTGAATCATCATCTTGGCTCACAGAAGTGC 161 DB 6581 TTTTCGACAGATGGGTAAGAACAGAGACTTGAATCATCATCTTGGCTCACAGAAGTGC 6640

QY	162	TTGAACAGAGAAAGTCATTAAGCGAAGCATACACGAACCTACGGCGAAGACCTTCTGAA	221
Db	6641	TGAACAGAGAAAGTCATTAAGCGAAGCATACACGAACCTACGGCGAAGACCTTCTGAA	6700
QY	222	GACTCTTGAACAAGAGGCTCTCAAGATTGG-----	253
Db	6701	GACTCTTGAACAAGAGGCTCTCAAGATTGGAGTTTCGAGTTTCGACCTTAAAGTATC	6766
QY	254	-----AG	255
Db	6761	GACCTTTAAACATCTTGTGTGAAGTCATGTATCTCAATTGGTTTTTTGTGCATGAGACAG	6822
QY	256	AGAGCTATCTTGTGTGGACTCTTGAAACCCGGTGAAGCGATGCTTTATTTGGCTAATGAA	315
Db	6821	AGAGCTATCTTGTGTGGACTCTTGAAACCCGGTGAAGCGATGCTTTATTTGGCTAATGAA	6880
QY	316	GCTACAAAAGATGAGCTTCAAGCAACCAAGTTCTTAATGAAAGTTGCTGCAAGACGA	375
Db	6881	GCTACAAAAGATGAGCTTCAAGCAACCAAGTTCTTAATGAAAGTTGCTGCAAGACGA	6940
QY	376	TCAACGCAAGCTGCTTCAAGCTAAGCAAGCTTACCATGCTCGCTACAAAGAGCTCTTGAA	435
Db	6941	TCAACGCAAGCTGCTTCAAGCTAAGCAAGCTTACCATGCTCGCTACAAAGAGCTCTTGAA	7000
QY	436	GAGAGCGTTGCTCACCAACACTACCGGTGACTTCAGAAAAGCTTTGGTCTCTTGTAAC	495
Db	7001	GAGAGCGTTGCTCACCAACACTACCGGTGACTTCAGAAAAGCTTTGGTCTCTTGTAAC	7060
QY	496	TCATACAGGTACGAAGAGATGAAGTGAACATGATGGCTTACCAAGAAAGCTAAGCTG	555
Db	7061	TCATACAGGTACGAAGAGATGAAGTGAACATGATGGCTTACCAAGAAAGCTAAGCTG	7120
QY	556	GTCCATGAGAAAATCAAGACCAAGCACTACATGATGAGATGTATTAGATCTTGTC	615
Db	7121	GTCCATGAGAAAATCAAGACCAAGCACTACATGATGAGATGTATTAGATCTTGTC	7180
QY	616	ACAAGAAGCAAGCTCAATCAATGCTTAAACCGTACCAAGATGATCATGCGAG	675
Db	7181	ACAAGAAGCAAGCTCAATCAATGCTTAAACCGTACCAAGATGATCATGCGAG	7248
QY	676	GAAATTCGAAGAGCTTGAGGAAGAGATGATGATCAAGTCCCTGGCACTTTGAG	735
Db	7241	GAAATTCGAAGAGCTTGAGGAAGAGATGATGATCAAGTCCCTGGCACTTTGAG	7300
QY	736	TCAACCATTCAGTGGCTTGCAAGACCAAGAGCTTTACTTTGTCAGATGCTTCGTTCAACA	795
Db	7301	TCAACCATTCAGTGGCTTGCAAGACCAAGAGCTTTACTTTGTCAGATGCTTCGTTCAACA	7360
QY	796	ATCAACCAAAACTGGAAGCTGATGAAGAGCACTCACTAAGATTGTGACCAACAAGAGCTGAG	855
Db	7361	ATCAACCAAAACTGGAAGCTGATGAAGAGCACTCACTAAGATTGTGACCAACAAGAGCTGAG	7420
QY	856	ATTGACTTGAAGGTCATTGGAAGAGATACCAAGCGCAGAAACAGATTCCTTTGAGACAA	915
Db	7421	ATTGACTTGAAGGTCATTGGAAGAGATACCAAGCGCAGAAACAGATTCCTTTGAGACAA	7480
QY	916	GCTATTTACCAAGACACTCGTGAAGATTACGAGAAGATGCTCGTGCACCTTCGCGTGA	975
Db	7481	GCTATTTACCAAGACACTCGTGAAGATTACGAGAAGATGCTCGTGCACCTTCGCGTGA	7540
QY	976	GATGATGCTTAATCAATCAATCTTCACAGAGAAAATTAAGCTGCTCAAGCTCTTGTT	1035
Db	7541	GATGATGCTTAATCAATCAATCTTCACAGAGAAAATTAAGCTGCTCAAGCTCTTGTT	7600
QY	1036	ATTCCTTAATCCTCCTCTCTCTCTCTTTGATGAGTTCAAACTGTTGATTTTGTGTTCTAC	1095
Db	7601	ATTCCTTAATCCTCCTCTCTCTCTCTTTGATGAGTTCAAACTGTTGATTTTGTGTTCTAC	7660
QY	1096	AAAAACCTTGTGTTCTGTGTGTGTGTTGAGTTCCTAATATATGCAAAAGAGAGAGA	1155
Db	7661	AAAAACCTTGTGTTCTGTGTGTGTGTTGAGTTCCTAATATATGCAAAAGAGAGAGA	7720
QY	1156	CAGAGAGAACAGTGTGCTCTTAAAGTTATATATATATGAAGACATGGCCTAAAAA	1215

Db	7721	CAGAGAGAACGATGCTCTTAAAGTTATATATGAAGCAATTTGGCTAAAAACA	7780
Qy	1216	AAAAAAAAAA 1226	
Db	7781	CAGACTAACAA 7791	
RESULT 8			
LOCUS	CQ806050	954 bp	DNA
DEFINITION	Sequence 2461 from Patent WO2004035798.		linear
ACCESSION	CQ806050		
VERSION	CQ806050.1	GI:47111659	
KEYWORDS			
SOURCE			
ORGANISM	Arabidopsis thaliana (thale cress)		
REFERENCE	1	Inze D., de Veylder L. and Vlieghe K.	
AUTHORS		Identification of novel e2f target genes and use thereof	
TITLE		Patent: WO 2004035798-A 2461 29-APR-2004;	
JOURNAL		CropDesign N.V. (BE)	
FEATURES			
source		Location/Qualifiers	
		1..954	
		/organism="Arabidopsis thaliana"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:3702"	
ORIGIN			
Query Match	77.6%;	Score 954;	DB 6;
Best Local Similarity	100.0%;	Pred. No. 8e-231;	
Matches	954;	Conservative 0;	Mismatches 0;
		Indels 0;	Gaps 0;
Qy	34	ATGGGACTCTTAAGTTCTGATTCGTTCCTCTCCCTCTGATGATGCTGACCAATTG	93
Db	1	ATGGGACTCTTAAGTTCTGATTCGTTCCTCTCTCTCTGATGATGCTGACCAATTG	60
Qy	94	AGAACCGCTTTTGAAGATGGGGATACGAAGAGACTTTGATCATATCTTTGGCTCAC	153
Db	61	AGAACCGCTTTTGAAGATGGGGATACGAAGAGACTTTGATCATATCTTTGGCTCAC	120
Qy	154	AGAGTGTCTGAACAGAGGAAGTATCATGAGCAACATACCAAGAACTTACGGGAAAGAC	213
Db	121	AGAGTGTCTGAACAGAGGAAGTATCATGAGCAACATACCAAGAACTTACGGGAAAGAC	180
Qy	214	CTTCTCAAGACTCTTGAACAGAGACTCTCTACGATTTGAGAGAGCTATCTTTGGTGG	273
Db	181	CTTCTCAAGACTCTTGAACAGAGACTCTCTACGATTTGAGAGAGCTATCTTTGGTGG	240
Qy	274	ACTCTTGAAACCGGTGAGCGTGATGCTTTATTTGGCTATGAGCTTACAAAAGATGACT	333
Db	241	ACTCTTGAAACCGGTGAGCGTGATGCTTTATTTGGCTATGAGCTTACAAAAGATGACT	300
Qy	334	TCAAGCAACCAAGTTCTTATGGAAGTTGCTTGCAACAAGACATCAAGCAGCTGCTCAC	393
Db	301	TCAAGCAACCAAGTTCTTATGGAAGTTGCTTGCAACAAGACATCAAGCAGCTGCTCAC	360
Qy	394	GCTAGGCAAGCTTACATGCTCGCTACAGAAAGTCTCTTGAAAGAGAGCTTTGCTCACAC	453
Db	361	GCTAGGCAAGCTTACATGCTCGCTACAGAAAGTCTCTTGAAAGAGAGCTTTGCTCACAC	420
Qy	454	ACTACCGGTGACTTCAAAAAGCTTTTGGTTCTTTGTTACCTCATACAGGTACGAAGA	513
Db	421	ACTACCGGTGACTTCAAAAAGCTTTTGGTTCTTTGTTACCTCATACAGGTACGAAGA	480
Qy	514	GATGAAGTGAACATGACATTTGGCTTACCAAGAAAGCTTAAGCGGCTCATGAGAAATCAAG	573
Db	481	GATGAAGTGAACATGACATTTGGCTTACCAAGAAAGCTTAAGCGGCTCATGAGAAATCAAG	540
Qy	574	GACCAAGCACTACATGATGAGATGTTATTAGATCTTGTTCACAAAGCAAGAAAGCTCAG	633

|||||  
Db 541 GACACGACCTACATGATGATGATGATTTATAGATCTTGTCCACAAAGAAAGCAAGCTCAG 600  
QY 634 ATCAATGCTACTTTTAAACGCTTACCAAGATGATCATGCGAAGAAATTTCTCAAGATCTT 693  
Db 601 ATCAATGCTACTTTTAAACGCTTACCAAGATGATCATGCGAAGAAATTTCTCAAGATCTT 660  
QY 694 GAGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 753  
Db 661 GAGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
QY 754 ACAAGACCAAGAGCTTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 813  
Db 721 ACAAGACCAAGAGCTTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
QY 814 GATGAAGAGACTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 873  
Db 781 GATGAAGAGACTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
QY 874 GAGAGAGAGTACAGAGCGCAGAGAACAGATCTCTTGGAGAAAGCTATTACCAAGACACT 933  
Db 841 GAGAGAGAGTACAGAGCGCAGAGAACAGATCTCTTGGAGAAAGCTATTACCAAGACACT 900  
QY 934 CGTGAAGATTACAGAAAGATGCTGTCGCACTTCTCGTGAAGATGATGCTTAA 987  
Db 901 CGTGAAGATTACAGAAAGATGCTGTCGCACTTCTCGTGAAGATGATGCTTAA 954  
RESULT 9 954 bp mRNA linear PLN 07-MAY-2003  
AF332435 LOCUS  
DEFINITION Arabidopsis thaliana clone C00123 (e) putative Ca2+-dependent  
ACCESSION AF332435 GI:12083277  
VERSION AF332435.1  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 954)  
REFERENCE  
AUTHORS Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L.,  
Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H.,  
Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shin,P., Southwick,A.,  
Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.  
Arabidopsis Open Reading Frame (ORF) Clones  
Unpublished  
2 (bases 1 to 954)  
REFERENCE  
AUTHORS Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Eguu,P.,  
Lee,J.M., Toriumi,M., Yu,G., Brooks,S., Chao,Q., Chen,H.,  
Karlin-Neumann,G., Kim,C., Iam,B., Miranda,M., Nguyen,M.,  
Palm,C.J., Shin,P., Southwick,A., Davis,R.W., Ecker,J.R. and  
Theologis,A.  
Direct Submission  
Submitted (28-DEC-2000) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
3 (bases 1 to 954)  
REFERENCE  
AUTHORS Yamada,K., Dale,J.M., Hsuan,V.W., Onodera,C.S., Quach,H.L.,  
Toriumi,M., Wong,C., Wu,H.C., Yu,G., Yuan,S., Chen,H.,  
Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shin,P., Southwick,A.,  
Tripp,M.G., Wu,T., Davis,R.W., Ecker,J.R. and Theologis,A.  
Direct Submission  
Submitted (07-MAY-2003) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
Annotation based on July 2002 version of the Arabidopsis genome  
submitted to Genbank.  
FEATURES  
source location/Qualifiers  
1..954  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="1"

/clone="C00123 (e)"  
/ecotype="Columbia"  
/note="This clone is now in pUNI 51 vector. Previously,  
it was in pUNI-T3-D/V5-His-10PO under the clone number  
C00123."  
1..954  
/gene="Atlg35720"  
1..954  
/gene="Atlg35720"  
CDS  
1..954  
/gene="Atlg35720"  
/codon\_start=1  
/evidence=experimental  
/product="putative Ca2+-dependent membrane-binding protein  
anexin"  
/protein\_id="AA048798.1"  
/db\_xref="GI:12083278"  
/translation="MATIKVDSVPAPSDDAEQLRTAREGKCTNEDLIISILAHSAE  
QKVIROAVHETYGEDLTKTDKELSNFERRAILMTLPBGRDMLLNEATKRTSS  
NOVMEVACTRTSTGLHARAYHARVYAEEDVAHHHTDPRLLVSLVTSYRYS  
DEVNMTLAKOEAQLVHEKIKDHYNDEVDVIRILSTRSAQINAEFRYVODHGEILK  
SLEBGDDDKFLALRSTICLREPLFEVDVLRSAINKTGDEBALTRIVYTRAIED  
LKVIQEVQRNRSIFLEKAITKDRRGDYERKMLVALLGRDA"  
ORIGIN  
Query Match 77.6%; Score 954; DB 8; Length 954;  
Best Local Similarity 100.0%; Pred. No. 8e-231;  
Matches 954; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 34 ATGGGACCTCTTAAGTTTCTGATTTCTGTTCTGCTCTCTCTCTGATGATGCTGAGCAATTG 93  
Db 1 ATGGGACCTCTTAAGTTTCTGATTTCTGTTCTGCTCTCTCTCTGATGATGCTGAGCAATTG 60  
QY 94 ACAACCGCTTTGAAGATGGGTACGACGAGACTTGATCATATCTTGGCTCAC 153  
Db 61 AGAACCGCTTTGAAGATGGGTACGACGAGACTTGATCATATCTTGGCTCAC 120  
QY 154 AGAAGTGCTGAACAGAGAAAGTCAATCAGGCAAGCATACAGAAACCTACGCGCAAGAC 213  
Db 121 AGAAGTGCTGAACAGAGAAAGTCAATCAGGCAAGCATACAGAAACCTACGCGCAAGAC 180  
QY 214 CTTTCCAAAGCTTTGACAGAGAGCTCTTAAACGATTTGAGAGAGCTATCTTGTTCG 273  
Db 181 CTTTCCAAAGCTTTGACAGAGAGCTCTTAAACGATTTGAGAGAGCTATCTTGTTCG 240  
QY 274 ACTCTTGAACCGGAGAGGCTGATGCTTATTTGGCTAATGAAGTACAAAAGATGAGACT 333  
Db 241 ACTCTTGAACCGGAGAGGCTGATGCTTATTTGGCTAATGAAGTACAAAAGATGAGACT 300  
QY 334 TCAAGCAACCAAGTTCTTATGGAAGTGTCTTGCACAAGACATCAACGAGCTGCTTCA 393  
Db 301 TCAAGCAACCAAGTTCTTATGGAAGTGTCTTGCACAAGACATCAACGAGCTGCTTCA 360  
QY 394 GCTAGCAAGCTTACAGAGCTGCTGCAAGAAAGTCTCTTGAAGAGAGAGTGTCTACAC 453  
Db 361 GCTAGCAAGCTTACAGAGCTGCTGCAAGAAAGTCTCTTGAAGAGAGAGTGTCTACAC 420  
QY 454 ACTACCGGAGCTTACAGAAAGCTTTGGTTTCTCTTGTATCTCTATACAGGATCAAGAA 513  
Db 421 ACTACCGGAGCTTACAGAAAGCTTTGGTTTCTCTTGTATCTCTATACAGGATCAAGAA 480  
QY 514 GATGAAGTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 573  
Db 481 GATGAAGTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
QY 574 GACAAAGCTACTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 633  
Db 541 GACAAAGCTACTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
QY 634 ATCAATGCTACTTTTAAACGCTTACCAAGATGATCATGCGAAGAAATTTCTCAAGATCTT 693  
Db 601 ATCAATGCTACTTTTAAACGCTTACCAAGATGATCATGCGAAGAAATTTCTCAAGATCTT 660  
QY 694 GAGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 753



Db 661 GAGGAGGAGATGATGATGACAGTTCCTTGCACTTTTGAGTCAACATTCAGTGGCTTG 720  
Qy 754 ACAAGACCAAGCTTTACTTGTGATGATTTCTTGTTCAGCAATCAACAAACTGGAAC 813  
Db 721 ACAAGACCAAGCTTTACTTGTGATGATTTCTTGTTCAGCAATCAACAAACTGGAAC 780  
Qy 814 GATGAAGAGCACTCATAGATTTGTGACACAGAGCTGATTTGATTTGAGTCACT 873  
Db 781 GATGAAGAGCACTCATAGATTTGTGACACAGAGCTGATTTGATTTGAGTCACT 840  
Qy 874 GAGAGAGATTCACAGCGAGGAGCAACATTCCTTTGAGAAAGCTATTACCAAGACAT 933  
Db 841 GAGAGAGATTCACAGCGAGGAGCAACATTCCTTTGAGAAAGCTATTACCAAGACAT 900  
Qy 934 CGTGAAGATTCAGGAAAGATGCTGTCGCACTTTCTCGGTGAAGATGATGCTTAA 987  
Db 901 CGTGAAGATTCAGGAAAGATGCTGTCGCACTTTCTCGGTGAAGATGATGCTTAA 954

RESULT 10  
AY356355 954 bp mRNA linear PLN 01-SEP-2004  
LOCUS AY356355  
DEFINITION Brassica juncea annexin mRNA, complete cds.  
ACCESSION AY356355  
VERSION AY356355.1 GI:38049639  
KEYWORDS  
SOURCE  
ORGANISM Brassica juncea  
Brassica juncea  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; euroside II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 954)  
Jami,S.K. and Kirti,P.B.  
Amexin cDNA from the floral bud of Brassica juncea  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 954)  
AUTHORS Jami,S.K. and Kirti,P.B.  
TITLE Direct Submission  
AUTHORS  
TITLE Submitted (30-JUL-2003) Department of Plant Sciences, School of  
JOURNAL Life Sciences, University of Hyderabad, Gachibowli, Hyderabad,  
Andhra Pradesh 500046, India  
Location/Qualifiers  
1. .954  
FEATURES  
source  
/organism="Brassica juncea"  
/mol\_type="mRNA"  
/db\_xref="taxon:3707"  
/issue\_type="floral bud"  
1. .954  
/codon\_start=1  
/product="annexin"  
/protein\_id="AAK10457.1"  
/db\_xref="GI:38049640"  
/translation="MATLKVSIVSPSPEDAOLKSAFDMGTNEELIISILARSAE  
ORKLIRQTHESFGEDELKSEKELSDPRALIMTLFEGEDALLVNPATKRTSS  
NOVMEVACTRSTOLHARQAYAPKRSIEDVAHHTTGDRKLLVLSVSRVREG  
EYVNTLAKBEAKLIHEKIKQKHNDPFIIRIISTRKAQIMNTFNRYDPNRSBEIK  
SLEBGEDDKFLGLRSTIOCLTRBELFYDVRKSAIKNTGTDGEGALTRIVATRAEID  
LKVIQDEYQRNSIPIEKAITKDTGRDYERKMLIALGEEDA"

ORIGIN  
Query Match 64.3%; Score 790.8; DB 8; Length 954;  
Best Local Similarity 89.3%; Pred. No. 1.9e-189;  
Matches 852; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

Qy 34 ATGGCGACTTTAAGTTTCTGATTCCTTCTGCTCTCTGATGATGCGAGCAATG 93  
Db 1 ATGGCGACTTTAAGTTTCTGATTCCTTCTGCTCTCTGATGATGCGAGCAATG 60  
Qy 94 AGAACCGCTTTGAGGATGGGGTACAGACAGAGACTTATCATATCAATCTTGGCTAC 153  
Db 61 AAAAGCCCATTTGATGGATGGGGTACAGACAGAGAAATGATCATATCAATCTTGGCTAC 120  
Qy 154 AGAAGTCTGAACAGAGAAAGTCAATGACGAGCAATACAGAAACCTACGGGGAAGAC 213

Db 121 AGAAGTCTGAACAGAGAAAGTCAATGATCAGGCAACATACATGAAATCTTTGGAGAGAT 180  
Qy 214 CTTCTCAAGCTTTGACAGAGAGCTCTTACAGATTTGAGAGAGCTATCTTGTGGG 273  
Db 181 CTTCTTAAAGTCTTGAAGAGAACTTACAGAGAGCTCTTGAAGAGAGCTTGTCTGG 240  
Qy 274 ACTTTGAACCCGGTGAAGCTGATGCTTTATTTGGCTTATGAACCTACAAAGATGACT 333  
Db 241 ACTTTGAACCCGGTGAAGCTGATGCTTTATTTGGCTTATGAACCTACAAAGATGACT 300  
Qy 334 TCAAGCAACCAAGTTCTTATGAGATGCTTGAACAAGACATCAACGAGCTGCTTAC 393  
Db 301 TCAAGCAACCAAGTTCTTATGAGATGCTTGAACAAGACCTTACGAGCTTCTTAC 360  
Qy 394 GCTAAGCAACCTTACAGATGCTGCTTCAAGAAAGCTCTTGAAGAGAGCGTGTCAAC 453  
Db 361 GCTAAGCAACCTTACAGATGCTGCTTCAAGAAAGCTCTTGAAGAGAGATGCTCAAC 420  
Qy 454 ACTACCGGTGATTCAGAAAGCTTTGGTTCCTTGTATACCTCATACAGTACGAAAG 513  
Db 421 ACCACCGGTGATTCAGAAAGCTTTGGTTCCTTGTATACCTCATACAGTACGAAAG 480  
Qy 514 GATGAAGTGAACATGACATTTGCTTGAACAAGAGCTAAGCTGCTCAAGAAATCAAG 573  
Db 481 GAAGAGTGAACATGACATTTGCTTGAACAAGAGCTAAGCTGCTCAAGAAATCAAG 540  
Qy 574 GACAGCACTACATGATGAGATGATTTAGATTTGCTTGAACAAGAGAGAGCTCAG 633  
Db 541 GACAGCACTACATGATGAGATGATTTAGATTTGCTTGAACAAGAGAGAGAGCTCAG 600  
Qy 634 ATCAATGCTTACTTTTAAAGCTTACCAAGATGATGAGAGAGAGAAATTCAGAGTCTT 653  
Db 601 ATCAATGCTTACTTTTAAAGCTTACCAAGATGATGAGAGAGAGAAATTCAGAGTCTT 660  
Qy 694 GAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 753  
Db 661 GAGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
Qy 754 ACAAGACCAAGCTTTACTTGTGATGATTTCTTGTTCAGCAATCAACAAACTGGAAC 813  
Db 721 ACAAGACCAAGCTTTACTTGTGATGATTTCTTGTTCAGCAATCAACAAACTGGAAC 780  
Qy 814 GATGAAGAGCACTCATAGATTTGTGACACAGAGCTGATTTGATTTGAGTCACT 873  
Db 781 GATGAAGAGCACTCATAGATTTGTGACACAGAGCTGATTTGATTTGAGTCACT 840  
Qy 874 GAGAGAGATTCACAGCGAGGAGCAACATTCCTTTGAGAAAGCTATTACCAAGACAT 933  
Db 841 GAGAGAGATTCACAGCGAGGAGCAACATTCCTTTGAGAAAGCTATTACCAAGACAT 900  
Qy 934 CGTGAAGATTCAGGAAAGATGCTGTCGCACTTTCTCGGTGAAGATGATGCTTAA 987  
Db 901 CGTGAAGATTCAGGAAAGATGCTGTCGCACTTTCTCGGTGAAGATGATGCTTAA 954

RESULT 11  
GHU89609 1141 bp mRNA linear PLN 30-AUG-1998  
LOCUS GHU89609  
DEFINITION Gossypium hirsutum fiber annexin mRNA, complete cds.  
ACCESSION U89609  
VERSION U89609.1 GI:3493171  
KEYWORDS  
SOURCE  
ORGANISM Gossypium hirsutum (upland cotton)  
Gossypium hirsutum  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; euroside II; Malvales; Malvaceae; Malvoideae; Gossypium.  
1 (bases 1 to 1141)  
Shin,H. and Brown,R.M.Jr.  
Direct Submission  
JOURNAL Submitted (13-FEB-1997) Botany, The Univ. of Texas at Austin,  
Austin, TX 78713, USA









**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 03:05:29 ; Search time 805 Seconds  
(without alignments)  
9045.071 Million cell updates/sec

```
Title: US-10-690-564-1
Perfect score: 1230
Sequence: 1 ccacgcgtccgaacactaa.....aaaaaaaaaaaaaaaa 1230
```

Scoring table: IDENTITY\_NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

1: N\_GeneSeq\_16Dec04.\*  
2: GeneSeqn1980s.\*  
3: GeneSeqn1990s.\*  
4: GeneSeqn2000s.\*  
5: GeneSeqn2001as.\*  
6: GeneSeqn2001bs.\*  
7: GeneSeqn2002as.\*  
8: GeneSeqn2002bs.\*  
9: GeneSeqn2003as.\*  
10: GeneSeqn2003bs.\*  
11: GeneSeqn2003cs.\*  
12: GeneSeqn2004as.\*  
13: GeneSeqn2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1162.4	93.1	1192	3	AAC34418	Aac34418 Arabidopsis
2	1144.8	93.1	1122	3	AAC34549	Aac34549 Arabidopsis
3	1141.4	92.8	1182	3	AAC34339	Aac34339 Arabidopsis
4	954	77.6	954	12	ADJ63871	Adj63871 Plant 11
5	954	77.6	954	12	ADN74566	Adn74566 Thale c1
6	635	51.6	635	6	ABO66078	Abg66078 Arabidopsis
7	497.8	40.5	1184	13	ADRE3461	Adre3461 Cotton c
8	473.4	38.5	1058	10	ADB94854	Adb94854 Program
9	471.8	38.4	2588	3	AAA67100	Aaa67100 Eucalyptu
c 10	470	38.2	487	10	ADBE1994	Adbe1994 Arabidop
11	468.6	38.1	1293	3	AAA67128	Aaa67128 Pinus rac
12	468.6	38.1	1293	4	AAAF4756	Aaf4756 Amexim-1
13	468.6	38.1	1293	10	ADB94943	Adb94943 Program
14	457.2	37.2	1160	3	AAC46038	Aac46038 Arabidops
15	455.6	37.0	1156	3	AAC49728	Aac49728 Arabidops
16	453.2	36.8	954	12	ADN74780	Adn74780 Thale cr
17	449.6	36.6	1157	10	ADP7996	Adp7996 Synchrotr
18	426	34.6	951	3	AAC41736	Aac41736 Arabidops
19	424	34.5	1025	3	AAC49738	Aac49738 Arabidops
20	413.6	33.6	1027	3	AAC37764	Aac37764 Arabidops

## ALIGNMENTS

21	364.2	29.6	789	3	AAA67131	AA67131 Pinus rad
22	364.2	29.6	789	3	AAA67098	AA67098 Eucalyptu
23	364.2	29.6	789	4	AAFA4759	AAFA4759 Anmexin-1
24	364.2	29.6	789	10	ADB94456	ADB94456 Program
25	347.8	28.3	420	3	AAc37667	AAc37667 Arabidops
26	327.2	26.6	502	13	ACN57334	ACN57334 Cotton g
27	327	26.6	591	13	ACN60265	ACN60265 Cotton g
28	326	26.5	589	13	ACN58684	ACN58684 Cotton g
29	322.2	26.2	600	13	ACN60840	ACN60840 Cotton g
30	315.8	25.7	586	13	ACN60745	ACN60745 Cotton g
31	294.2	23.9	570	12	ADG07773	ADG07773 Polymucl
32	282.6	23.0	570	13	ACN58229	ACN58229 Cotton g
33	271	22.0	608	13	ACN53624	ACN53624 Cotton g
34	270.6	22.0	2132	3	AAAF7099	AAAF7099 Pinus rad
35	270.4	22.0	454	3	AAAC6827	AAAC6827 Arabidops
36	266.2	21.6	504	13	ADRG3464	ADRG3464 Cotton c
37	230.8	18.8	504	13	ACN59774	ACN59774 Cotton g
38	213.4	17.3	437	12	ADP95395	ADP95395 Cotton e
39	210	17.1	704	4	AAFA4760	AAFA4760 Anmexin-1
40	208.4	16.9	704	3	AAAF7132	AAAF7132 Pinus rad
41	208.4	16.9	704	10	ADP94947	ADP94947 Program
42	206.8	16.8	704	3	AAAF7096	AAAF7096 Pinus rad
43	198	16.1	396	12	ADP93369	ADP93369 Cotton e
44	186.2	15.1	484	3	AAAF7129	AAAF7129 Pinus rad
45	186.2	15.1	484	4	AAFA4757	AAFA4757 Anmexin-1

ALIGNMENTS

RESULT 1

AAAC34418

ID

AAAC34418 standard; DNA; 1192 BP.

XX

AAAC34418;

AC

17-OCT-2000 (first entry)

DT

XX

Arabidopsis thaliana DNA fragment SEQ ID NO: 6580.

DE

XX

Hybridisation assay; genetic mapping; gene expression control;

KM

protein identification; signal transduction pathway; metabolic pathway;

KW

promoter; termination sequence; ss.

XX

XX

Arabidopsis thaliana.

OS

PM

EP1033405-A2.

XX

PD

06-SEP-2000.

XX

PF

25-FEB-2000; 2000BP-00301439.

XX

XX

25-FEB-1999; 99US-0121625P.

PR

05-MAR-1999; 99US-0123180P.

PR

09-MAR-1999; 99US-0123548P.

PR

23-MAR-1999; 99US-0125788P.

PR

25-MAR-1999; 99US-0126264P.

PR

29-MAR-1999; 99US-0126785P.

PR

01-APR-1999; 99US-0127462P.

PR

06-APR-1999; 99US-0128234P.

PR

08-APR-1999; 99US-0128714P.

PR

16-APR-1999; 99US-0129645P.

PR

19-APR-1999; 99US-0130077P.

PR

21-APR-1999; 99US-0130449P.

PR

23-APR-1999; 99US-0130510P.

PR

28-APR-1999; 99US-0130691P.

PR

30-APR-1999; 99US-0131449P.

PR

04-MAY-1999; 99US-0132407P.

PR

05-MAY-1999; 99US-0132484P.

PR

06-MAY-1999; 99US-0132485P.

PR

06-MAY-1999; 99US-0132486P.

PR

PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0133256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137228P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145951P.  
PR 28-JUL-1999; 99US-0146368P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 06-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149428P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149903P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156566P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158368P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.

```
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      94.5%; Score 1162.4; DB 3; Length 1192;
Best Local Similarity 99.1%; Pred. No. 1,2e-286;
Matches 1169; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 10 CGAAGACTTAAAGTGAAGAAAAATGCGACTTTAGGGTTCTGATTCTGTCTGGCT 69
DB 12 CAAAACATRAAGTGAAGAAAAATGCGACTCTTAAGTTCTGATTCTGTCTGGCT 71
QY 70 CCTTCTGATGATGCTGAGCAATTGAAACCGCTTTTGAAGATGGGTAAGAACGAGAC 129
DB 72 CCTTCTGATGATGCTGAGCAATTGAAACCGCTTTTGAAGATGGGTAAGAACGAGAC 131
QY 130 TTGATATATTCATCTTTGGCTCAGAAAGTGTGAACAGAGAAAGTCAACAGGACGA 189
DB 132 TTGATATATTCATCTTTGGCTCAGAAAGTGTGAACAGAGAAAGTCAACAGGACGA 191
QY 190 TACCAAGAAACCTAACGCGGAGAACTTTCAAGACTCTTGAAGAGGCTCTTAAGAT 249
DB 192 TACCAAGAAACCTAACGCGGAGAACTTTCAAGACTCTTGAAGAGGCTCTTAAAGAC 251
QY 250 TTGAGAGAGCTATCTTTGTGTGACTCTTGAACCGGATGAGGCTTATTTATTTGGCT 309
DB 252 TTGAGAGAGCTATCTTTGTGTGACTCTTGAACCGGATGAGGCTTATTTATTTGGCT 311
QY 310 AATGAAGCTAACAAAAAGATGATTTCAAGCAACCAAGTTCTTATGGAAGTTGCTTGA 369
DB 312 AATGAAGCTAACAAAAAGATGATTTCAAGCAACCAAGTTCTTATGGAAGTTGCTTGA 371
QY 370 AGGACATCAACGAGCTGCTTCAAGCAAGCTTACAGTCTGCTCAACAAGAGCT 429
DB 372 AGGACATCAACGAGCTGCTTCAAGCAAGCTTACAGTCTGCTCAACAAGAGCT 431
QY 430 CTGGAAGAGAGCGTTGCTCAACCACTACCGGTGACTTGAAGGCTTTTGTGTTCTCT 489
DB 432 CTGGAAGAGAGCGTTGCTCAACCACTACCGGTGACTTGAAGGCTTTTGTGTTCTCT 491
QY 490 GTTAACTCTTACAGGATGCAAGAGAGATGAAGTGAACATGACATTTGGCTTAAAGAGCT 549
DB 492 GTTAACTCTTACAGGATGCAAGAGAGATGAAGTGAACATGACATTTGGCTTAAAGAGCT 551
QY 550 AAGCTGCTCATAGGAAAAATCAAGGCAAGCACTAACAATGATGAGATGTTATTGAATC 609
DB 552 AAGCTGCTCATAGGAAAAATCAAGGCAAGCACTAACAATGATGAGATGTTATTGAATC 611
QY 610 TTGTCCACAAGAGCAAGAGCTGATCAATGCTTATTAACCGTTACCAAGATGATCAT 669
DB 612 TTGTCCACAAGAGCAAGAGCTGATCAATGCTTATTAACCGTTACCAAGATGATCAT 671
QY 670 GGGGAGAAATTTCTCAAGAGCTTTGAGAGAGAGATGATGACAAAGTTCTTGGACATT 729
DB 672 GGGGAGAAATTTCTCAAGAGCTTTGAGAGAGAGATGATGACAAAGTTCTTGGACATT 731
QY 730 TTGAGGTCAACCATTCAGTGTCTTGAACAAGCCAGACTTTTATCTTGTGATCTTGGT 789
DB 732 TTGAGGTCAACCATTCAGTGTCTTGAACAAGCCAGACTTTTATCTTGTGATCTTGGT 791
QY 790 TCAGCAATCAACAAAATGGAATCTGATGGAAGAGCACTGATGAATTTGACCAACAGA 849
DB 792 TCAGCAATCAACAAAATGGAATCTGATGGAAGAGCACTGATGAATTTGACCAACAGA 851
```

```
QY 850 GCTGAGATTGACTTGAAGGTCAATTGAGAGAGTACACCGCAGGAAACGATTCCTTTG 909
DB 852 GCTGAGATTGACTTGAAGGTCAATTGAGAGAGTACACCGCAGGAAACGATTCCTTTG 911
QY 910 GAGAAAGCTATTACCAAGACACCTGAGAGATTACGAGAAAGATGCTGTCGACATTCTC 969
DB 912 GAGAAAGCTATTACCAAGACACCTGAGAGATTACGAGAAAGATGCTGTCGACATTCTC 971
QY 970 GGTGAAGATGATGCTTAATCAATCAATCTCCACAGAGAAACATAGCTGCTCAAGCT 1029
DB 972 GGTGAAGATGATGCTTAATCAATCAATCTCCACAGAGAAACATAGCTGCTCAAGCT 1031
QY 1030 TCTGTATCTCTTAATCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1089
DB 1032 TCTGTATCTCTTAATCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1091
QY 1090 TTCTACAAAAACCTTGTGTTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1149
DB 1092 TTCTACAAAAACCTTGTGTTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1151
QY 1150 GAGAGACAGAGAAACCAAGTGTGTCTCTTAAGTTATATA 1189
DB 1152 GAGAGACAGAGAAACCAAGTGTGTCTCTTAAGTTATATA 1191

RESULT 2
AAC34549
ID AAC34549 standard; DNA; 1212 BP.
XX
AC AAC34549;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 7037.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
XX
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
```

PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134376P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135153P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138647P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139563P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 14-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.



PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140659P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 18-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.



PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 92.8%; Score 1141.4; DB 3; Length 1182;  
 Best Local Similarity 98.6%; Pred. No. 2.7e-281;  
 Matches 1163; Conservative 0; Mismatches 11; Indels 5; Gaps 1;

QY 10 CGAAACCTAAAGTAGAAGAAAAAGCCGCTTTAAGGTTCTGATCTGTTCTGCT 69  
 DB 9 CAAAACCTAAAGTAGAAGAAAAAGCCGCTTTAAGGTTCTGATCTGTTCTGCT 68  
 QY 70 CTTTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGCGGTGAGAACGAGAC 129  
 DB 69 CTTTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGCGGTGAGAACGAGAC 123  
 QY 130 TTGATCATATCAATCTTGGCTCAAGAGTCTGAAAGAGAAAGTCTAAGGCAAGCA 189  
 DB 124 TTGATCATATCAATCTTGGCTCAAGAGTCTGAAAGAGAAAGTCTAAGGCAAGCA 183  
 QY 190 TACCAAGAACTTAAGCGGCAAGACCTTTCTGAAGCTCTTGAAGAGCTCTTAAGAT 249  
 DB 184 TACCATGAAACTTAAGCGGCAAGACCTTTCTGAAGCTCTTGAAGAGCTCTTAAGAT 243  
 QY 250 TTGAGAGAGCTATCTTGTGTGGAAGCTTGAAGCCGCTGAGGTGATGCTTATTTGGCT 309  
 DB 244 TTGAGAGAGCTATCTTGTGTGGAAGCTTGAAGCCGCTGAGGTGATGCTTATTTGGCT 303  
 QY 310 AATGAGCTCAAAAAGATGAGACTTCAAGCAACCAAGTTCTTATGAAAGTTGCTTGACA 369  
 DB 304 AATGAGCTCAAAAAGATGAGACTTCAAGCAACCAAGTTCTTATGAAAGTTGCTTGACA 363  
 QY 370 AGGACATCAACGAGCTGCTTCAAGCTGAGCAAGCTTACCATGCTGCTCAAGAAAGTCT 429  
 DB 364 AGGACATCAACGAGCTGCTTCAAGCTGAGCAAGCTTACCATGCTGCTCAAGAAAGTCT 423  
 QY 430 CTTGAAAGAGAGCTTGTGCTACACACTACCGGTGACTTGAAAGCTTTTGGTTTCTT 489  
 DB 424 CTTGAAAGAGAGCTTGTGCTACACACTACCGGTGACTTGAAAGCTTTTGGTTTCTT 483  
 QY 490 GTTACCTCATACAGGTACGAAAGAGATGAGTGAACATGATGCTGCTAAGCAAGAGCT 549  
 DB 484 GTTACCTCATACAGGTACGAAAGAGATGAGTGAACATGATGCTGCTAAGCAAGAGCT 543  
 QY 550 AAGCTGCTCATGAGAAATCAAGAGCAAGCACTACATGATGAGATGTTATTAGATC 609  
 DB 544 AAGCTGCTCATGAGAAATCAAGAGCAAGCACTACATGATGAGATGTTATTAGATC 603  
 QY 610 TTGTCCACAAAGCAAGCTCAGATCATGCTTAAACGCTTACCAAGATGATCAT 669  
 DB 604 TTGTCCACAAAGCAAGCTCAGATCATGCTTAAACGCTTACCAAGATGATCAT 663  
 QY 670 GGGGAGAAATTTCTCAAGAGCTTTGAGAGAGATGATGATGACAAAGTCTTGAAGCT 729  
 DB 664 GGGGAGAAATTTCTCAAGAGCTTTGAGAGAGATGATGATGACAAAGTCTTGAAGCT 723  
 QY 730 TTGAGGTCAACCATGAGTCTTGAAGAGCAAGCTTAACTTTGATGATGTTCTTCTG 789  
 DB 724 TTGAGGTCAACCATGAGTCTTGAAGAGCAAGCTTAACTTTGATGATGTTCTTCTG 783  
 QY 790 TCGCAATCAAAAGCTGAGTGAAGAGAGCACTCTGTAATTTGACCAAGCA 849  
 DB 784 TCGCAATCAAAAGCTGAGTGAAGAGAGCACTCTGTAATTTGACCAAGCA 843  
 QY 850 GCTGAGTTGACTTGAAGTCAATGAGAGAGTCAAGGCGAGAAAGCAAGCTTCCCTTG 969  
 DB 844 GCTGAGTTGACTTGAAGTCAATGAGAGAGTCAAGGCGAGAAAGCAAGCTTCCCTTG 903  
 QY 910 GAGAAAGCTATTAACAAAGCACTGCTGAGATTAAGAGAAAGTCTGCTGCACTTCTC 969  
 DB 904 GAGAAAGCTATTAACAAAGCACTGCTGAGATTAAGAGAAAGTCTGCTGCACTTCTC 963

QY 970 GGTGAAGATGAGTCTTAATCAATCATCTCCCAAGAGAAACAATAGCTGCTTACAGCT 1029  
 DB 964 GGTGAAGATGAGTCTTAATCAATCATCTCCCAAGAGAAACATTAAGCTGCTTACAGCT 1023  
 QY 1030 TCTGTTATCTTATCTCCCTCTCTCTCTTGTGATGAGTTCGTTGATTTTGT 1089  
 DB 1024 TCTGTTATCTTATCTCCCTCTCTCTCTTGTGATGAGTTCGTTGATTTTGT 1083  
 QY 1090 TTCTACAAAACCTTGTGTTGTTCTGTGTGTTGATTCCTAATATACAAAGA 1149  
 DB 1084 TTCTACAAAACCTTGTGTTGTTCTGTGTGTTGATTCCTAATATACAAAGA 1143  
 QY 1150 GAGAGACAGAGAGAACAGTGTGCTCTTAAGTTAT 1188  
 DB 1144 GAGAGACAGAGAGAACAGTGTGCTCTTAAGTTAT 1182

RESULT 4  
 ADJ63871  
 ID ADJ63871 standard; DNA; 954 BP.

AC ADJ63871;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX

DE Plant lipid metabolism protein OO-10 gene SEQ ID NO:73.  
 XX  
 XX de; gene; plant; lipid metabolism protein; LMP; seed storage compound;  
 KM transgenic plant.  
 XX  
 OS Unidentified.

Key Location/Qualifiers  
 FT 1..954  
 FT /\*tag= a  
 FT /product= "OO-10"  
 XX

PN WO2004013304-A2.  
 XX

PD 12-FEB-2004.  
 XX

PF 04-AUG-2003; 2003WO-US024364.  
 XX

PR 02-AUG-2002; 2002US-040803P.  
 XX

PA (BAD1) BASF PLANT SCI GMBH.  
 XX

PI Miltendorf V, Haertel HA, Bauer J, Oswald O;  
 XX

DR WPI; 2004-157121/15.  
 XX

DR P-PSDB; ADJ63872.  
 XX

PT New lipid metabolism proteins and nucleic acids, useful in producing  
 PT transgenic plants with increased levels of seed storage compound, e.g.  
 PT lipid, a fatty acid, a starch or a seed storage protein.  
 XX

XX Claim 1; SEQ ID NO 73; 115bp; English.  
 XX

XX The invention relates to novel isolated lipid metabolism proteins (LMP)  
 CC and encoding nucleic acids comprising a polynucleotide sequence encoding  
 CC a polypeptide that functions as a modulator of seed storage compounds in  
 CC a plant. The LMP nucleic acid is useful in producing transgenic plants  
 CC with increased levels of seed storage compound, e.g. lipid, a fatty acid,  
 CC a starch or a seed storage protein, as markers for specific regions of  
 CC the genome and for evolutionary and protein structural studies. The  
 CC present sequence represents an LMP nucleic acid of the invention.

XX Sequence 954 BP; 288 A; 198 C; 229 G; 239 T; 0 U; 0 Other;

Query Match 77.6%; Score 954; DB 12; Length 954;  
 Best Local Similarity 100.0%; Pred. No. 2e-233;  
 Matches 954; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 34 ATGGGACCTCTTAAGGTTCTGATTCCTGTTCTGTTCTGTTGATGAGCAATTG 93
   |||||
DB 1 ATGGGACCTCTTAAGGTTCTGATTCCTGTTCTGTTCTGTTGATGAGCAATTG 60
QY 94 AGAACCGCTTTTGAAGATGGGGTACGAAAGAGACTTATCATATCAATCTTGGCTCAC 153
   |||||
DB 61 AGAACCGCTTTTGAAGATGGGGTACGAAAGAGACTTATCATATCAATCTTGGCTCAC 120
QY 154 AGAAGTCTGAAAGAGAAAGTATCATAGGCAAGCATACCAAGAACTTACGGCCGAAGAC 213
   |||||
DB 121 AGAAGTCTGAAAGAGAAAGTATCATAGGCAAGCATACCAAGAACTTACGGCCGAAGAC 180
QY 214 CTTCTCAAGACTTTTGAAGAGAGCTCTGTAAGATTTGAGAGAGACTTCTTGTGTGG 273
   |||||
DB 181 CTTCTCAAGACTTTTGAAGAGAGCTCTGTAAGATTTGAGAGAGACTTCTTGTGTGG 240
QY 274 ACTCTTGAACCCGGTGAAGGCTGATGCTTATTTGGCTTATGAAAGCTACAAAAGATGACT 333
   |||||
DB 241 ACTCTTGAACCCGGTGAAGGCTGATGCTTATTTGGCTTATGAAAGCTACAAAAGATGACT 300
QY 334 TCAGCAACCAAGTTCTTATGAAAGTTGCTTGCAAGAGCATCAACGAGCTGCTTAC 393
   |||||
DB 301 TCAGCAACCAAGTTCTTATGAAAGTTGCTTGCAAGAGCATCAACGAGCTGCTTAC 360
QY 394 GCTAGGCAAGCTTACGATGCTGCTACAAAGAGCTCTGAAAGAGAGCGTGTGCAACAC 453
   |||||
DB 361 GCTAGGCAAGCTTACGATGCTGCTACAAAGAGCTCTGAAAGAGAGCGTGTGCAACAC 420
QY 454 ACTACCGGTGACTTCAGAAAGCTTTTGGTTTCTTCTTGTATACCTATACAGATACGAAGA 513
   |||||
DB 421 ACTACCGGTGACTTCAGAAAGCTTTTGGTTTCTTCTTGTATACCTATACAGATACGAAGA 480
QY 514 GATGAAGTGAACATGACATTTGGCTTACGAAGAAAGCTTGTCCATGAGAAATCAAG 573
   |||||
DB 481 GATGAAGTGAACATGACATTTGGCTTACGAAGAAAGCTTGTCCATGAGAAATCAAG 540
QY 574 GACAAGCACTACATGATGAGATGATTTATGAACTTTGGCCCAAGAAAGCAAAAGCTCAG 633
   |||||
DB 541 GACAAGCACTACATGATGAGATGATTTATGAACTTTGGCCCAAGAAAGCAAAAGCTCAG 600
QY 634 ATCAATGCTACTTTTAAACGTTTACCAAGATGATCATGGGAGGAAATTTCTCAAGACTTT 693
   |||||
DB 601 ATCAATGCTACTTTTAAACGTTTACCAAGATGATCATGGGAGGAAATTTCTCAAGACTTT 660
QY 694 GAGGAAGAGATGATGATGACAAAGTTCTTGGCACTTTTGAAGTCAACCATTCAGTCTTG 753
   |||||
DB 661 GAGGAAGAGATGATGATGACAAAGTTCTTGGCACTTTTGAAGTCAACCATTCAGTCTTG 720
QY 754 ACAAGACCAAGAGCTTACTTGTGAGATGTTCTTGTTCAAGCAATCAACAAACCTGGAAC 813
   |||||
DB 721 ACAAGACCAAGAGCTTACTTGTGAGATGTTCTTGTTCAAGCAATCAACAAACCTGGAAC 780
QY 814 GATGAAGAGCACTCACTAAGATTTGACCAAGAGCTGAGATGATGATGAGTCAAT 873
   |||||
DB 781 GATGAAGAGCACTCACTAAGATTTGACCAAGAGCTGAGATGATGATGAGTCAAT 840
QY 874 GGAAGAGAGTACCAAGGCGGAGAAACAGCATTTCTTTGAGAAAGCTTATTCAAAGCACT 933
   |||||
DB 841 GGAAGAGAGTACCAAGGCGGAGAAACAGCATTTCTTTGAGAAAGCTTATTCAAAGCACT 900
QY 934 CGTGAAGATTACGAAGAAGATGCTGCTGCACTTCTGGTGAAGATGATGCTTAA 987
   |||||
DB 901 CGTGAAGATTACGAAGAAGATGCTGCTGCACTTCTGGTGAAGATGATGCTTAA 954
```

```
RESULT 5
ADN74566 standard; cDNA; 954 BP.
XX
AC ADN74566;
XX
DT 15-JUL-2004 (first entry)
```

```
XX DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants Segid 2461.
XX KM gene; se; plant; transgenic; E2Fa/Dpa transcription factor;
XX KW growth regulator; animal feed product; thale cress;
XX KM cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX OS Arabidopsis thaliana.
XX PN WO2004035798-A2.
XX PD 29-APR-2004.
XX PF 20-OCT-2003; 2003WO-EP011658.
XX PR 18-OCT-2002; 2002EP-00079408.
XX (CROP-) CROPDISEIGN NV.
XX Inze D, De Veylder L, Vlieghe K;
XX WPI; 2004-348466/32.
XX DR P-PSDB; ADN74567.
XX PT Altering plant characteristics, useful for producing plants for enzyme or
XX PT pharmaceutical production comprises modifying in a plant, expression of
XX PT one or more nucleic acids and/or modifying level or activity of one or
XX PT more proteins.
XX Claim 1; SEQ ID NO 2461; 134pp; English.
XX PS
XX CC This invention relates to a novel method for altering one or more plant
XX CC characteristics. Specifically, it refers to identifying genes that are up
XX CC - or down-regulated in transgenic plants overexpressing the heterodimeric
XX CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
XX CC alter plant characteristics accordingly. The present invention describes
XX CC generating transgenic plants for the production of growth regulators,
XX CC enzymes, therapeutic, pharmaceuticals and animal feed products, where
XX CC the altered plant characteristics are selected from increased yield or
XX CC biomass, enhanced survival capacity, stress tolerance, plant architecture
XX CC or physiology, altered endoreduplication, biochemistry, signal
XX CC transduction, storage lipid mobilisation and/or altered photosynthesis,
XX CC each relative to the corresponding wild type plants. Accordingly, these
XX CC sequences can also be useful as positive or negative selectable markers
XX CC during transformation of cells or tissues. The identified genes play a
XX CC role in a variety of biological processes such as DNA replication, cell
XX CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as
XX CC transcription factors. This polynucleotide sequence is thale cress cDNA
XX CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa
XX CC transcription factor, given in an exemplification of the invention.
XX SQ Sequence 954 BP; 288 A; 198 C; 223 G; 233 T; 0 U; 0 Other;
XX
XX Query Match 77.6%; Score 954; DB 12; Length 954;
XX Best Local Similarity 100.0%; Pred. No. 2e-233;
XX Matches 954; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 ATGGGACCTTTAAGTTCGATTCCTGTTCTGTTCTGTTGATGAGCAATTG 93
   |||||
DB 1 ATGGGACCTTTAAGTTCGATTCCTGTTCTGTTCTGTTGATGAGCAATTG 60
QY 94 AGAACCGCTTTTGAAGATGGGGTACGAAAGAGACTTATCATATCAATCTTGGCTCAC 153
   |||||
DB 61 AGAACCGCTTTTGAAGATGGGGTACGAAAGAGACTTATCATATCAATCTTGGCTCAC 120
QY 154 AGAAGTCTGAAAGAGAAAGTATCATAGGCAAGCATACCAAGAACTTACGGCCGAAGAC 213
   |||||
DB 121 AGAAGTCTGAAAGAGAAAGTATCATAGGCAAGCATACCAAGAACTTACGGCCGAAGAC 180
QY 214 CTTCTCAAGACTTTTGAAGAGAGCTCTGTAAGATTTGAGAGAGACTTCTTGTGTGG 273
   |||||
DB 181 CTTCTCAAGACTTTTGAAGAGAGCTCTGTAAGATTTGAGAGAGACTTCTTGTGTGG 240
```

QY 274 ACTTGACCCGGTGAGCGGATGCTTTATTTGGCTAATGAAGCTACAAAAGTGGACT 333  
DB 241 ACTCTTAAACCCGGTGAGCGGATGCTTTATTTGGCTAATGAAGCTACAAAAGTGGACT 300  
QY 334 TCAGCAACCAAGTTCTTATGAGAGTTGCTTGACAAAGACATCAAGCAGCTGCTTAC 393  
DB 301 TCAAGCAACCAAGTTCTTATGAGAGTTGCTTGACAAAGACATCAAGCAGCTGCTTAC 360  
QY 394 GCTTAGCAAGCTTACATGCTCGCTACAGAAAGTCTCTTGAAGAGAGCTTCTCACCAC 453  
DB 361 GCTTAGCAAGCTTACATGCTCGCTACAGAAAGTCTCTTGAAGAGAGCTTCTCACCAC 420  
QY 454 ACTACCGGTGACTTCAAGAAAGCTTTGGTTCTCTTGTACCTCATCAGGTGCGAAGA 513  
DB 421 ACTACCGGTGACTTCAAGAAAGCTTTGGTTCTCTTGTACCTCATCAGGTGCGAAGA 480  
QY 514 GATGAAGTGAACATGATGCTGCTTACCAAGAAAGCTAAGTGTGCTCATGAGAAATCAG 573  
DB 481 GATGAAGTGAACATGATGCTGCTTACCAAGAAAGCTAAGTGTGCTCATGAGAAATCAG 540  
QY 574 GACAAAGCACTAATGATGAGATGTTATTTGAATCTTGTCCACAGAAAGCTCAG 633  
DB 541 GACAAAGCACTAATGATGAGATGTTATTTGAATCTTGTCCACAGAAAGCTCAG 600  
QY 634 ATCAATGCTACTTTTAACCGTTACCAAGATGATGCGAGAAATTCACAGAGCTT 693  
DB 601 ATCAATGCTACTTTTAACCGTTACCAAGATGATGCGAGAAATTCACAGAGCTT 660  
QY 694 GAGGAAGAGATGATGATGACAAAGTCTTGCACTTTTGAAGTCAACATTCAGTCTTG 753  
DB 661 GAGGAAGAGATGATGATGACAAAGTCTTGCACTTTTGAAGTCAACATTCAGTCTTG 720  
QY 754 ACAAGACCAAGCTTTACTTTGCTGATGTTCTTGTACGATCAACAAAGCTGAACT 813  
DB 721 ACAAGACCAAGCTTTACTTTGCTGATGTTCTTGTACGATCAACAAAGCTGAACT 780  
QY 814 GATGAAGAGCACTCACTAATTTGATCCCAAGAGCTGAGATGATGCTTGAAGTCAAT 873  
DB 781 GATGAAGAGCACTCACTAATTTGATCCCAAGAGCTGAGATGATGCTTGAAGTCAAT 840  
QY 874 GGAAGAGATGATGATGACAAAGTCTTGCACTTTTGAAGAAAGTATTTACCAAGACT 933  
DB 841 GGAAGAGATGATGATGACAAAGTCTTGCACTTTTGAAGAAAGTATTTACCAAGACT 900  
QY 934 CGTGAAGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 987  
DB 901 CGTGAAGATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954

RESULT 6  
AB066078 ID AB066078 standard; DNA; 635 BP.  
XX  
XX AB066078;  
XX  
DT 21-AUG-2002 (first entry)  
XX  
DE Arabidopsis thaliana polynucleotide SEQ ID NO 655.  
XX  
XX Arabidopsis thaliana; thale cress; plant; transgenic; GMO; disease;  
KM stress; metabolic pathway; biosynthetic pathway; nutrition; fungicide;  
KM insecticide; antibiotic; ds.  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
PN US2002059663-A1.  
XX  
PD 16-MAY-2002.  
XX  
XX 26-JAN-2001; 2001US-00770149.  
XX  
XX 27-JAN-2000; 2000US-0178506P.  
XX

PA (GORL/) GORLACH J.  
PA (ANY/) AN Y.  
PA (HAM/) HAMILTON C M.  
PA (PRIC/) PRICE J L.  
PA (RAIN/) RAINES T M.  
PA (YUY/) YU Y.  
PA (RAME/) RAMEKA J G.  
PA (PAGE/) PAGE A.  
PA (MATH/) MATHEW A V.  
PA (LEDF/) LEDFORD B L.  
PA (WOES/) WOESNER J P.  
PA (HAAS/) HAAS W D.  
PA (GARC/) GARCIA C A.  
PA (KRICK/) KRICKER M.  
PA (SLAT/) SLATER T.  
PA (DAVI/) DAVIS K R.  
PA (ALLE/) ALLEN K.  
PA (HOFF/) HOFFMAN N.  
PA (HURB/) HURBAN P.  
XX  
PI Gorlach J, An Y, Hamilton CM, Price JI, Raines TM, Yu Y,  
PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD,  
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N,  
PI Hurban P;  
XX  
DR WPI; 2002-479224/51.  
XX  
PT New nucleic acid that hybridizes to Arabidopsis thaliana sequences,  
PT useful e.g. for preparing transgenic plants with increased resistance or  
PT altered metabolism.  
XX  
PS Claim 1; SEQ ID NO 655; 40pp + Sequence Listing; English.  
XX  
CC The invention relates to nucleic acids (i) that hybridise under stringent  
CC conditions to any of 999 sequences (AB065424-AB066422) or their  
CC fragments, (ii) are used to express the corresponding polypeptides (ii) or  
CC to produce genetically modified plant cells or transgenic plants, which  
CC may have improved resistance to disease or stress, or altered  
CC metabolic/biosynthetic pathways (for production of commercial,  
CC nutritional or medicinal products), or generally any trait of interest,  
CC or can be used to screen for biologically active agents (e.g. fungicides,  
CC insecticides and antibiotics). Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from the USPTO at  
CC [seqdata.uspto.gov/sequence.html?DocID=99909770149](http://seqdata.uspto.gov/sequence.html?DocID=99909770149)  
XX  
SQ Sequence 635 BP; 197 A; 133 C; 147 G; 158 T; 0 U; 0 Other;  
XX

Query Match 51.6%; Score 635; DB 6; Length 635;  
Best Local Similarity 100.0%; Pred. No. 5.8e-152;  
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 16 ACTTAAAGTAGAAGAAATGGGCACTTTAAGTTTGTGATTTCTGCTGCTGCTTCT 75  
DB 1 ACTTAAAGTAGAAGAAATGGGCACTTTAAGTTTGTGATTTCTGCTGCTGCTTCT 60  
QY 76 GATGATGCTGAGCAATTTGAAGAACCGCTTTTGAAGATGGGGTACGAAGAGACTTGAATC 135  
DB 61 GATGATGCTGAGCAATTTGAAGAACCGCTTTTGAAGATGGGGTACGAAGAGACTTGAATC 120  
QY 136 ATATCAATCTTGGCTCAAGAAAGTGTGAACAGAGAAAGTATTCAGGCAAGATACAC 195  
DB 121 ATATCAATCTTGGCTCAAGAAAGTGTGAACAGAGAAAGTATTCAGGCAAGATACAC 180  
QY 196 GAAACCTAGCGGCAAGACCTTCTCAAGACTTTTGAAGAGAGCTCTTAAGATTTGAG 255  
DB 181 GAAACCTAGCGGCAAGACCTTCTCAAGACTTTTGAAGAGAGCTCTTAAGATTTGAG 240  
QY 256 AGAGCTATCTTGTGTGAGCTTTGAACCCGGTGAACGGTATGCTTATTTGGCTAATGA 315  
DB 241 AGAGCTATCTTGTGTGAGCTTTGAACCCGGTGAACGGTATGCTTATTTGGCTAATGA 300  
QY 316 GCTTCAAAAAGATGAGCTTCAAGCAACCAAGTTCTTATGAAAGTCTTGCAAGAGCA 375

DB 301 GGTACAAAAGATGAGCTTCAAGCAACCAAGTTCTTAGAAGTCTTGACACAGACA 360  
QY 376 TCACGCGAGCTGCTTACCGCTAGGCAAGCTTACCATGCTGCTACAGAAAGTCTTTGAA 435  
DB 361 TCACCCAGAGCTGCTTACCGCTAGGCAAGCTTACCATGCTGCTACAGAAAGTCTTTGAA 420  
QY 436 GAGAGAGTGTGCTACGACCACTACCGGTGACTTACGAAAGCTTTGTTCTCTTTTACC 495  
DB 421 GAGAGAGTGTGCTACGACCACTACCGGTGACTTACGAAAGCTTTGTTCTCTTTTACC 480  
QY 496 TCATACAGGTACGAGAGAGATGAAGTGAACATGACATTTGGCTTAAGCAAGAGCTAAGCTG 555  
DB 481 TCATACAGGTACGAGAGAGATGAAGTGAACATGACATTTGGCTTAAGCAAGAGCTAAGCTG 540  
QY 556 GTCCATGAGAAATTCAGAGCAAGCACTACATGATGAGATGTTATTAGAAATCTTGTCC 615  
DB 541 GTCCATGAGAAATTCAGAGCAAGCACTACATGATGAGATGTTATTAGAAATCTTGTCC 600  
QY 616 ACAAGAGCAAGAGCTGATCATGCTACTTTTAA 650  
DB 601 ACAAGAGCAAGAGCTGATCATGCTACTTTTAA 635  
RESULT 7  
ADRe63461  
ID ADRe63461 standard; cDNA; 1184 BP.  
XX  
AC ADRe63461;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Cotton cDNA sequence, SEQ ID 4242.  
XX  
KM Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;  
KM drought tolerance; plant disease resistance; galactomannan; lignin;  
KM plant growth regulator; heat tolerance; herbicide tolerance;  
KM homologous recombination; extreme osmotic condition tolerance;  
KM pathogen resistance; pest resistance; yield; photosynthesis; seed oil;  
KM stress resistance.  
XX  
OS Gossypium hirsutum.  
OS  
PN US2004181830-A1.  
XX  
PD 16-SEP-2004;  
XX  
PF 29-JAN-2004 2004US-00767795.  
XX  
PR 07-MAY-2001 2001US-00849529.  
PR 12-DEC-2001 2001US-00021323.  
XX  
PA (KOVA/) KOVALIC D K.  
PA (ZHOU/) ZHOU Y.  
PA (CAOY/) CAO Y.  
XX  
PI Kovalic DK, Zhou Y, Cao Y;  
PI  
DR WPI; 2004-667718/65.  
XX  
PT New recombinant nucleic acid molecules and polypeptides from Gossypium  
PT hirsutum, useful for producing plants with improved biological  
PT characteristics (e.g. improved plant cold or drought tolerance).  
XX  
XX Claim 1; SEQ ID NO 4242; 14pp; English.  
XX  
XX The invention relates to a recombinant polynucleotide comprising any of  
CC the 58798 Cotton plant cDNA sequences mentioned in the specification.  
CC Also a recombinant polypeptide comprising any of the 58798 amino acid  
CC sequences mentioned in the specification and producing a plant having an  
CC improved property. Producing a plant having an improved property  
CC comprises transforming a plant with a recombinant construct comprising a  
CC promoter region functional in a plant cell operably joined to a

CC polynucleotide comprising a coding sequence for a polypeptide associated  
CC with the property, and growing the transformed plant. The polypeptide is  
CC useful for improving plant cold tolerance, manipulating growth rate in  
CC plant cells by modification of the cell cycle pathway, improving plant  
CC drought tolerance, providing increased resistance to plant disease,  
CC producing galactomannan (or lignin or plant growth regulators), improving  
CC plant heat tolerance, improving plant tolerance to herbicides, increasing  
CC the rate of homologous recombination in plants, improving plant tolerance  
CC to extreme osmotic conditions or to pathogens or pests, improving yield  
CC by modification of photosynthesis, modifying seed oil or protein yield  
CC and/or content, improving yield by modification of carbohydrate, nitrogen  
CC or phosphorus use and/or uptake, or improving yield by providing improved  
CC plant growth and development under at least one stress condition. The  
CC polynucleotide and polypeptide may also be used in recombinant DNA  
CC constructs, in physical arrays of molecules, as plant breeding markers,  
CC or in computer-based storage and analysis systems. The present sequence  
CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC [seqdata.uspto.gov/sequence.html?DocID=20040181830](http://seqdata.uspto.gov/sequence.html?DocID=20040181830). However only 6585  
CC polynucleotide sequences were available, the remaining 5213  
CC polynucleotides and all 58798 protein sequences were not present.  
XX  
SQ Sequence 1184 BP; 375 A; 233 C; 281 G; 295 T; 0 U; 0 Other;

Query Match 40.5%; Score 497.8; DB 13; Length 1184;  
Best Local Similarity 70.7%; Pred. No. 8:5e-117;  
Matches 691; Conservative 0; Mismatches 282; Indels 4; Gaps 2;

QY 12 AAACCTMAAGTAGAAGA-AAATGGCGACTCTTAAGGTTCTGATTCCTGCTGC 70  
DB 47 AAGAAATMAAGAGAGAGAGATGGCCACTTACAGTCCCAAGAGTTCCTTGG 106  
QY 71 CTTTGATGATGCTGAGCAATGAGAACCGCTTTTGAAAGATGGGTGCGAAGAGCT 130  
DB 107 TGTCTGAAGATTTGTAAACGCTTAAAGAAAGCTTTTCAGAGATGGAGTAATGAGGCT 166  
QY 131 TGATATATCAATCTTGGCTCACAGAGTGTGGAAGAGAAATCATCAGCAAGCAT 190  
DB 167 TATATATGATATATTTGGTCAAGAAATGCCAGACAGAAATCTGATGAAAAACCT 226  
QY 191 ACCAGAAACCTAGCGCGAAGACCTTCTCAAGACTCTTGACAGAGAGCTCTTAAAGATT 250  
DB 227 ACGTGAAGACCTATGAGAGAGATCTCTCAAGGCACTAGACAAAGAGCTCTGAATGACT 286  
QY 251 TCGAGAGAGCTATCTGTTGGAGCTCTTGAACCCGCGAGCGTATGTTATGCGCTA 310  
DB 287 TTGAGAGGCTGTTTGTCTTGGGCTTTGATCTGCTGAAGCTATGCTCTTTGGCTA 346  
QY 311 ATGAAGCTACAAAAGATGGACTTCAGAGCAACCAAGTCTTATGGAAGTGTGCTGCACA 370  
DB 347 ATGAAGCCACCAAAAGTGGAGCTTCAAGCAATCAAGTCTTATGGAATAGCTTGACAA 406  
QY 371 GGAATATCAACGAGCTGCTTCAAGCTAGGCAAGCTTACCATGCTGCTACAAAGATCTC 430  
DB 407 GGTCTGCCAACCAAGCTTCAAGCAAGCGAGGCTTATCATGCTGTTATGAAGTGGC 466  
QY 431 TTGAAGAGAGCTGCTGCTACCACTACCGGTGACTTCAGAAAGCTTTTGGTTCTCTTG 490  
DB 467 TTGAAGAGAGTGTGCTCATCACAGACTGGGAGCTTCGTAAGTCTCTCACTCTAG 526  
QY 491 TTACCTATACAGGTACGAGAGAGATGAAGTGAACATGATGCTAGCTAAGCAAGAGCTA 550  
DB 527 TGAATTCATACGATATGAGGAGAGAGTGAACATGAATTGGCGAAGAAAGAGGCGA 586  
QY 551 AGCTGTCATGAGAAATCAAGCAAGCACTACATGATGAGAGATGTTATTAGAAATCT 610  
DB 587 AGTGGCTTATGAGAAATTTTCAAGCAAGCTTACAGTATGATGATGATGATGATGATG 646  
QY 611 TGTCCAGAGAGCAAGAGTCAAGTCAATGCTTATTAACGCTTACCAAGATGATGATG 670  
DB 647 TGGCTACAGAGAGCAAGGACAGATCAATGCAACTGTGAATCATACAAAAAATGAATAG 706

QY 671 GCGAGAAATTCCTCAAGAGCTTGAAGAGAGATGATGATGACAGTTCCTGACCTTT 730  
 Db 707 GAATGACATTAACAGAGACTT---GAGGCTGATCTTAAGATGAGATTCCTTGACACTAC 763  
 QY 731 TGAAGTCAACCAATTCAGTCTTGACAAAGACCAAGCTTAACTTTGATGATGCTTTCGTT 790  
 Db 764 TAAGGTCCACAGTGAAGTGTGCTTATCCGAAAAGTATTTTGAAGAGTTCCTTGCC 823  
 QY 791 CAGAAATCAACAAACCTGGAATGATGATGAGAGACATCTAGATTTGACCAAGAG 850  
 Db 824 TAGCAATCAATAGACGAGAACCGATGATGATGAGAGCTTACTAGATTTGCACTAGGG 883  
 QY 851 CTGAGATTTGACTTGAAGGATCATTTGAGAGAGTACGACGCGAGAGACAGATTCCTTGG 910  
 Db 884 CTGAGGTTGATCTTAATATCATAGACATGATGATCCAGCAAGGAGACAGTCTCCACTGA 943  
 QY 911 AGAAAGCTATTACCAAGACACTGCGAGATTAACGAGAAAGTCTGCGCACTTTCG 970  
 Db 944 CTGCGCTATTGTCAGAGACACTCATGAGACATGAAAAATTGCTGCTGTACTTGCAG 1003  
 QY 971 GTGAGATGATGCTTAA 987  
 Db 1004 GACATGTGAGAAATTGA 1020

RESULT 8  
 ADB94854  
 ID ADB94854 standard; cDNA; 1058 BP.  
 AC ADB94854;  
 DT 04-DEC-2003 (first entry)  
 DE Programmed cell death pathway protein annexin cDNA #6.  
 XX  
 XX Programmed cell death; plant development; plant cell cycle; ATL2; DAD1;  
 KM Dnaase; 11s; 16d; nucellin-like aspartic protease; annexin; prohibitin;  
 KM fan-like protein; race; retinoblastoma-related protein; SIVA; TRID;  
 KM TBGT; xylogenic kinase; pur-alpha; cyteine protease; RPP5-like protein;  
 KM gp 91 NADPH oxidase subunit; NPR-like protein; BAG-1;  
 KM defender against cell death; lethal leaf spot; lesion stimulating death;  
 KM seven in absentia; transcription initiation factor;  
 KM testis enhanced gene transcript; gene; ss.  
 XX  
 OS Eucalyptus grandis.  
 PN US2003082724-A1.  
 PD 01-MAY-2003.  
 PF 14-AUG-2002; 2002US-00219220.  
 PR 04-JUN-1999; 99US-00325932.  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 XX Plim B, Lasham A;  
 XX MPI; 2003-786916/74.  
 DR P-PSDB; ADB94896.  
 XX  
 PT New isolated polynucleotide useful for modulating programmed cell death,  
 PT altering the development cycle of plant cells, and subsequently modifying  
 PT plant development.  
 XX  
 PS Claim 1; Page 138; 214pp; English.  
 XX  
 CC The invention describes an isolated polynucleotide (1) comprising a  
 CC nucleotide sequence that is one of 145 fully defined sequences of 221-  
 CC 315 base pairs (bp), given in the specification, and/or its complements,  
 CC reverse complements, reverse sequences, or sequences having 75, 90 or 95  
 CC % sequence identity to or that hybridise under stringent hybridisation  
 CC conditions to one of the 145 sequences. The methods and compositions of

CC the present invention to do with (1) are useful for modulating programmed  
 CC cell death and thereby altering the development cycle of plant cells, and  
 CC altering plant development. This sequence encodes a protein associated  
 CC with the programmed cell death pathway.  
 XX  
 SQ Sequence 1058 BP; 306 A; 227 C; 274 G; 251 T; 0 U; 0 Other;  
 Query Match 38.5%; Score 473.4; DB 10; Length 1058;  
 Best Local Similarity 69.0%; Pred. No. 1,4e-110;  
 Matches 664; Conservative 0; Mismatches 296; Indels 3; Gaps 1;  
 QY 25 AGAAGAAAAATGCGACTCTTAAGGTTTCTGATTTCTGTCCTGCTCTTGAATGATGCT 84  
 Db 63 AGAAGAGAAATGCGACTCTTAAGGTTTCTGATTTCTGTCCTGCTCTTGAATGATGCT 122  
 QY 85 GAGCAATTGAGAACCGCTTTTGAAGATGGGATGAGCAAGAGAGACTTGAATATCAATC 144  
 Db 123 GAGCACTCCAAAAGCTTTGCGAGATGGGAGCAAGATGATGATCATATCATAT 182  
 QY 145 TTGGCTTACAGAAAGTCTGAACAGAGAAAGTATATAGGCAAGATACCGAAACTTAC 204  
 Db 183 CTGGCTACAGAAACGAGCGACGCGAAAGTATCCGAAACATATGCGGAGATAT 242  
 QY 205 GCGAAGACCTTCTCAAGACTCTTGAAGAGAGCTCTTAAGATTTGAGAGAGCTATC 264  
 Db 243 GCGAAGATCTTCTCAAGAGCTTGAAGAGAGCTCTTAAGATTTGAGAGATCTG 302  
 QY 265 TTGTTGTGACTTTGAACCCGCTGAGCGTATGCTTTATTTGGCTATGAGATCAAAA 324  
 Db 303 CTTCGTGAGCCCTGATCTGCGAGCGGATGATGATCTTGTCCATGATGATGATCAAG 362  
 QY 325 AGATGACTTCAAGCAACCAAGTCTTATGAGAAAGTCTTGAAGAGATCAAGCGAG 384  
 Db 363 AGATGACTTCAAGCAACCAAGTCTTATGAGAAAGTCTTGAAGAGATCAAGCGAG 422  
 QY 385 CTGCTTCAAGGAGAGAGCTTCAAGAGCTTCAAGAGAGCTTGAAGAGAGCTT 444  
 Db 423 TTATTCATGATGAGAGAGCTTCAAGAGCTTCAAGAGAGCTTGAAGAGAGCTT 482  
 QY 445 GCTCAACCACTACCGGATCTTCAAGAAAGCTTTGATTTCTTGTATCTTCACTATCAGG 504  
 Db 483 GCATATCACACTATCGGGGATTTCCGCAAGCTGCTTCTCTGCAAGTATCTTTCGG 542  
 QY 505 TACGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 564  
 Db 543 TATGAGGAGGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602  
 QY 565 AAAATCAAG 624  
 Db 603 AAGATTCAG 662  
 QY 625 AAGCTCAGATCAATGCTTCTTAAACCTTACCAAGATGATGATGATGATGATGATGAT 684  
 Db 663 AAGCTCAGATCAATGCTTCTTAAACCTTACCAAGATGATGATGATGATGATGATGAT 722  
 QY 685 AAGCTTGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 744  
 Db 723 AAGATCT---GAGGCTGATCAAAATGATGATGATGATGATGATGATGATGATGAT 779  
 QY 745 CAGTCTTGAAG 804  
 Db 780 AAGTCTTGAAG 839  
 QY 805 ACTGAACTGATGAG 864  
 Db 840 CTGGAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 899  
 QY 865 AAGCTATTTGAG 924  
 Db 900 CAGCGATTAAG 959  
 QY 925 AAAAGACTCGTGAAGATGAG 984

Db 1446 GCTCCGACAGCCTTCGCAGATGGGACAAATGAGAAGCTGATCATATCCATATTGGG 1505

genetically modified organism; transgenic organism; plant;

KW inhibitor testing; activator testing; modifier testing; fungicide;



KM insecticide; genetic function; genetic regulation; cellular metabolism;  
KM gene; ss.  
XX Arabidopsis thaliana.  
XX US2003115639-A1.  
XX 19-JUN-2003.  
XX 26-JAN-2001; 2001US-00770961.  
XX 27-JAN-2000; 2000US-0178466P.  
XX  
XX (GORL/) GORLACH J.  
PA (ANY/) AN Y.  
PA (HAM/) HAMILTON C M.  
PA (PRIC/) PRICE J L.  
PA (RAIN/) RAINES T M.  
PA (YUY/) YU Y.  
PA (RAME/) RAMEKA J G.  
PA (PAGE/) PAGE A.  
PA (MATH/) MATHEW A V.  
PA (LED/) LEDFORD B L.  
PA (WOES/) WOESSNER J P.  
PA (HAAS/) HAAS W D.  
PA (GARC/) GARCIA C A.  
PA (KRIC/) KRICKER M.  
PA (SLAT/) SLATER T.  
PA (DAVI/) DAVIS K R.  
PA (ALLE/) ALLEN K.  
PA (HOFF/) HOFFMAN N.  
PA (HURB/) HURBAN P.  
XX  
PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y,  
PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
PI Hurban P;  
XX  
XX WPI; 2003-810930/76.  
XX  
XX  
XX Novel Arabidopsis thaliana nucleic acids useful for generating  
PT genetically modified transgenic organisms, for screening biologically  
PT active agents such as fungicides, insecticides.  
XX  
XX Claim 1; SEQ ID NO 765; 44bp; English.  
XX  
XX The invention describes a nucleic acid (I) comprising a sequence capable  
CC of hybridizing under stringent conditions to any one of 999 fully defined  
CC Arabidopsis thaliana sequences (1) as given in specification e.g., 360,  
CC 1137, 455, 219, 472 nucleotides, etc, or its fragment. (I) is useful as a  
CC hybridization probe to complementary molecules in a cDNA library. (I) is  
CC also useful for generating genetically modified and transgenic organisms,  
CC usually plant cells and plants. A protein encoded by (I) is useful in  
CC screening assays to determine the effect of candidate inhibitors,  
CC activators or modifiers of the gene product. The protein is also useful  
CC for screening biologically active agents e.g., fungicides and  
CC insecticides. A genetically modified cell, comprising an exogenous  
CC nucleic acid, where the nucleic acid comprises transcription regulatory  
CC sequences operably linked to a sequence capable of hybridizing under  
CC stringent conditions to (1) is useful in the study of genetic function  
CC and regulation, for alteration of the cellular metabolism and for  
CC screening compounds that may affect the biological function of the gene  
CC or gene product. This sequence represents an Arabidopsis thaliana  
CC polynucleotide of the invention.  
XX  
XX Sequence 487 BP; 136 A; 96 C; 88 G; 167 T; 0 U; 0 Other;  
XX  
XX Query Match 38.2%; Score 470; DB 10; Length 487;  
XX Best Local Similarity 100.0%; Pred. No. 7.8e-110;  
XX Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 761 CAGAGCTTACTTGTGATGTTCTTGGTTCAGCAATCAAAAAGTGAAGTGAAG 820  
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 487 CAGAGCTTACTTGTGATGTTCTTGGTTCAGCAATCAAAAAGTGAAGTGAAG 428  
Qy 821 GAGCACTCACTAAGATTGTGACCAAGAAGCTGAAGTTGACTTGAAGTCAATTGGAGAG 880  
Db 427 GAGCACTCACTAAGATTGTGACCAAGAAGCTGAAGTTGACTTGAAGTCAATTGGAGAG 368  
Qy 881 AGTACGAGGCGAGAAACAGATTCTTTGGAGAAAGCTATTACCAAGAAGTCAATTGGAGAG 940  
Db 367 AGTACGAGGCGAGAAACAGATTCTTTGGAGAAAGCTATTACCAAGAAGTCAATTGGAGAG 308  
Qy 941 ATTACGAGAAAGATGCTGTCGACACTTCTGCTGAAGATGCTTAATCAATCAATCTTC 1000  
Db 307 ATTACGAGAAAGATGCTGTCGACACTTCTGCTGAAGATGCTTAATCAATCAATCTTC 248  
Qy 1001 CACAGAGAAACATTAAGCTGCTGACAGCTTCTGTTATCTTATCTTCCTCTCTCTCT 1060  
Db 247 CACAGAGAAACATTAAGCTGCTGACAGCTTCTGTTATCTTATCTTCCTCTCTCTCT 188  
Qy 1061 TTGATGAGTTTCAAAATGTTTGAATTTTGTCTTCAAAAACCTTGTGTTGTTGTTG 1120  
Db 187 TTGATGAGTTTCAAAATGTTTGAATTTTGTCTTCAAAAACCTTGTGTTGTTGTTG 128  
Qy 1121 TGTGTTGAGTTCTTAATTAATGAAGAAAGAGAGACAGAGAACCAAGTGTGTTCTTA 1180  
Db 127 TGTGTTGAGTTCTTAATTAATGAAGAAAGAGAGACAGAGAACCAAGTGTGTTCTTA 68  
Qy 1181 AGTTATATATATATATGAAGACATTTGGCTTAAAAAATTTTTTTTTTTTTTTT 1230  
Db 67 AGTTATATATATATATGAAGACATTTGGCTTAAAAAATTTTTTTTTTTTTTTT 18  
XX  
XX RESULT 11  
XX AAA67128  
XX ID AAA67128 standard; DNA; 1293 BP.  
XX AC AAA67128;  
XX  
XX 31-OCT-2000 (first entry)  
XX  
XX Pinus radiata annexin nucleotide sequence SEQ ID NO:109.  
XX  
XX Eucalyptus grandis; pinus radiata; Monterey pine; modification;  
KM plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;  
XX transgenic plant; ds.  
XX  
XX Pinus radiata.  
XX OS  
XX PN WC200022092-A2.  
XX PD 20-APR-2000.  
XX PF 08-OCT-1999; 99WC-NZ000169.  
XX PR 13-OCT-1998; 98US-00170862.  
XX PR 11-AUG-1999; 99US-0148426P.  
XX  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
XX  
XX Blokeberg LN;  
XX  
XX WPI; 2000-339328/29.  
XX P-PDB; AAB16321.  
XX  
XX New genes encoding proteins involved in a plant polysaccharide  
PT biosynthetic pathway, useful for modulating or altering the  
PT polysaccharide content, composition or structure of the plant.  
XX  
XX Claim 1; Page 93; 301bp; English.  
XX  
XX The present invention describes isolated polynucleotides (PN) comprising  
CC a sequence selected from one of 835 nucleotide sequences given in  
CC AAA67073 to AAA67907, their (reverse) complements, sequences producing an

CC Expectation (E) value of 0.01 or less compared to the 835 sequences,  
 CC sequences at least 50% identical to them. 200, 100, 40 or 20-mers of the  
 CC 835 sequences or sequences that are degenerately equivalent or allelic to  
 CC the 835 sequences. The polynucleotides are used to modify the activity of  
 CC a polypeptide involved in a polysaccharide biosynthetic pathway in the  
 CC plant. They are especially used to modulate or alter the polysaccharide  
 CC content, composition or structure of the plant. AAB16268 to AAB16340 are  
 CC proteins encoded by some of the polynucleotide sequence given in the  
 CC present invention

XX Sequence 1293 BP, 359 A, 290 C, 313 G, 331 T, 0 U, 0 Other;

Query Match 38.1%; Score 468.6; DB 3; Length 1293;

Best Local Similarity 68.4%; Pred. No. 2.5e-109;

Matches 664; Conservative 0; Mismatches 304; Indels 3; Gaps 1;

QY 30 AAAAATGGGCGACTTTAAGGTTTCTGATTTCTGCTCTCTTCTGATGATGCTGAGCA 89  
 DB 91 AACAAATGTCGACTCTCACCGTCCGACGACCTGCCCTGTAGCCGATGACTGCGAGCA 150  
 QY 90 ATTGAGAACCGCTTTGAAGATGGGGTACGAAGAGGACTTGATGATATCATCTTGGC 149  
 DB 151 GCTCCGAGAGCCTTGGCAGGATGGGAAACAATGAAAGCTGATCATATTCATTTGG 210  
 QY 150 TCACAGAAAGTGTGAACAGAGAAAGTCAATCAGGCAAGCATACCAAGAACTTACGGCGA 209  
 DB 211 TCATAGGAATGGCGCGAGAGAGAGTGAATGGCAAACTATGCGGAGCTTACGGCGA 270  
 QY 210 AGACCTTCTCAGACTCTTGACAGAGACTCTTAAAGATTTGAGAGACTATCTTGT 269  
 DB 271 GGACCTCTCAGAGCATTTGACAGAGAACTTACCAATGATTTGAGAGGCTGTGCTCT 330  
 QY 270 GTGAGCTCTTGAACCGGAGGAGTGATGCTTATTTGGCTAAAGAGCTACAAAGAGT 329  
 DB 331 TTGCTCATCTTGATCGGCTGAACGTATGCTGCTTGGCAATGAAGCGAGAAAGATG 390  
 QY 330 GACTTCAAGCAACCAAGTTCTTATGAAAGTTGCTTGACAGAGCATCAACGAGCTGCT 389  
 DB 391 GACTTCAAGCAACAGGTTCTCATGAAATAGCTGACAGAGGCTCTCCGAGCAAGTTGCT 450  
 QY 390 TCACGCTAGGCAAGCTTACATGCTGCTACAGAAAGTCTCTTGAAGAGGACTTGTCTCA 449  
 DB 451 TATGGCAAGCAAGCATATCATGCCGATCAAGAAAGTCAATGAAAGAGGACCTCCCTCA 510  
 QY 450 CCACACTACCGGAGCTTGCAAGAGCTTTGGTTCTCTTGTAACTCATACGATGATGCA 509  
 DB 511 CCAACAGACTGGAGATTTCTGTAAGTTGCTGTGATCTTGGAGGCTCCCTACGTAATGA 570  
 QY 510 AGGAGATGAAGTGAACATGACATTTGGCTAAGCAAGAGCTTACGCTGCTCATGAGAAAT 569  
 DB 571 TGGAGATGAGGTGAATATGACTTTGGAAAGCAGAGGCTTAAATCTCCAGAGAAAGAT 630  
 QY 570 CAAGGACAGCACTACATGATGAGATGTTATTTAGATTTTGTCCACAGAGCAAGC 629  
 DB 631 CTCAGGAAGGCTTATGCGCATAGGATCTCATAGAGATTTTGGCTACTATGAGCAAGC 690  
 QY 630 TCAGATCAATGCTACTTTTAAACGTTACCAAGATGATCAATGAGGAGAAATTTCTCAAG 689  
 DB 691 ACAGGTCATGCTACGCTGATATCATCAAAATGAGTTGAAATGATATCAACAAGGA 750  
 QY 690 TCTTGAAGAGAGATGATGATGACAAGTTCTTGGCACTTTGAGGTCAACCATTCAGTG 749  
 DB 751 TTT---GAAAACGATCCAAAAGACGGTTCCTTACTATATCTGAGCTTACATTAAGTG 807  
 QY 750 CTTGACAGACGACAGAGCTTTACTTTGTGATGTTCTTGTCTTCAAGCAATCAAAAATG 809  
 DB 808 CTTGACCTCGCTGAGAGATTTTGAAGTTCTTCTGCTAGCATCAATTAAGGAGAG 867  
 QY 810 AACTGATGAAGGAGCACTCACTGAATTTGACCAAGAGCTGAGATTTGACTTGAAGGT 869  
 DB 868 AACAGATGAAGGGGCTTGAACAGAGATTTGCTACACAGGGCCGAGGTTGACATGAAGTT 927  
 QY 870 CATTTGAGAGAGTACCAAGCGCAGGAACAGACTTCTTTGAGAAAGCTTATCAAGA 929

DB 928 TATTAATGAGAGTACACAGAGGAAATGATCTCTTCATGCTGATGTCAGAGCA 987  
 QY 930 CACTGTGAGATTTAGAGAAATGCTGCTCCGACTTCTCGGTGAAGATGATGCTTAATC 989  
 DB 988 CACTACTGAGACTATGAAAAATGCTTGTGATGATTTGGCCACGTCGAGGCTTGATT 1047  
 QY 990 AATCATCTCTC 1000  
 DB 1048 TACAACTACTC 1058

RESULT 12

AAP44756

ID AAP44756 standard; cDNA; 1293 BP.

XX AAP44756;

XX 27-MAR-2001 (first entry)

XX Annexin-like protein coding sequence #1.

XX Cell death modulator; programmed cell death; PCD; apoptosis;

XX forestry plant; ss.

XX Pinus radiata.

XX WO200075331-A1.

XX 14-DEC-2000.

XX 02-JUN-2000; 2000WO-NZ000086.

XX 04-JUN-1999; 99US-00325932.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX Flinn B, Lasham A;

XX WPI, 2001-061724/07.

XX P-PSDB; AAB65730.

XX Novel defender against cell death polynucleotide useful for modulating

XX programmed cell death pathway and specific development pathways in

XX forestry plant.

XX Claim 1; Page 54-55; 142pp; English.

XX The present invention relates to coding sequences (see AAP44740-F44840

XX and AAP44843-F44844) and proteins (see AAB65714-B65814) involved in

XX programmed cell death (PCD; apoptosis). The coding sequences and proteins

XX of the present invention are useful for modulating a PCD or cell death

XX pathway and various developmental pathways in a forestry plant, by stably

XX incorporating one of the present coding sequences into the genome of a

XX forestry plant, where the coding sequence provides a PCD pathway that is

XX not present in a native form of the forestry plant

XX Sequence 1293 BP, 359 A, 290 C, 313 G, 331 T, 0 U, 0 Other;

XX Query Match 38.1%; Score 468.6; DB 4; Length 1293;

XX Best Local Similarity 68.4%; Pred. No. 2.5e-109;

XX Matches 664; Conservative 0; Mismatches 304; Indels 3; Gaps 1;

Db	211	TCATGGAATGGGGCGCAGAGGAACTGATTCGGCAAAACTATATGCGGAGACTTACGGGGA	270
Qy	210	AGACCTTCTCAAGACTCTTGAACAAAGAGCTCTTAACGATTTCCAGAGACTATCTTGT	269
Db	271	GGACCTCCTCAAGGCAATGGACAGAGAACTTACCAATGATTTCCAGAGGCTGGTGGTCT	330
Qy	270	GTGCACTCTTGAACCCGGTAGCGGTGATGCTTTATTTGGCTAATGAAGCTACAAAAAGAT	329
Db	331	TTGGTCACCTTGAATCCGGGCTGAACGTGATGCGTAATTTGGCGAATGAAGCGACGAAAAAGAT	390
Qy	330	GACCTTCAAGCAACCAAGTTCTTATGGAAGTCTTCCACAAAGACATCAAGCAGCTGCT	389
Db	391	GACCTTCAAGCAACCAAGTTCTTATGGAAGTCTTCCACAAAGACATCAAGCAGCTGCT	450
Qy	390	TCAACCTTAGGCAAGCTTACCATGCTCGCTCAAAAGATCTCTTGAAGGAGAGTGGCTCA	449
Db	451	TATGGCAAGACAAAGCATATCATGCCCGATACAAAGATCATGGAAGAGAGCTGGCTCA	510
Qy	450	CCACACTACCGGTGACTTCAGAAAGCTTTGGTTTCTCTGTATTCATACAGGTACGA	509
Db	511	CCACAACACTGGAGATTTTCGTAATTTGGTGGTACCTCTTGGGAGCTCTACCGTAATGA	570
Qy	510	AGGAATGAAGTGAACATGACATTTGGCTTAAGCAAGAGCTAAGCTGGTTCATGAGAAAT	569
Db	571	TGGAATGAGGTGAATATGACTTTTGGCAAAAGCAAGAGCTAAGATATCCACAGAGAT	630
Qy	570	CAAGGACAAGCACTCAATGATGAGATGTTTATTAAGATCTTGTCCACAAGAGCAAG	629
Db	631	CTCAGAGAGAGGCTTATGGCCATGAGAGATCTCTAATGAATTTTGGCTTACGAGCAAGC	690
Qy	630	TCAGATCAATGCTACTTTTAAACCGTTACCAAGATGATGATGGCGAGAAATTTCTCAAG	689
Db	691	ACAGGTCAATGCTATGCTGAAATCACTACAAAATAGATTTGSAATGATATCAACAGGA	750
Qy	690	TCTTGAAGGAGAGATGATGATGACAAAGTTCTTGCACCTTTAGAGTCAACCATTCAGTG	749
Db	751	TTT---GAAACTGATCCAAAAGACGCGTCTCTACTATACGAGCTACGTAAGTG	807
Qy	750	CTTGAACAAAGCCAGGCTTTACTTGTGAGATGTTTGGTTCAGCAATCAACAAACATGG	809
Db	808	CTGAGCTGCCCTTGAAAGATTTTGAAGAGTTCTTGTGCTAGCCATCAATTAACGAGG	867
Qy	810	AACTGATGAAGAGGACCTCATAGAAATTTGTGACCAAGAGCTGAGATTTGACTTGAAGT	869
Db	868	AAACAATGAAGGGGCTCTGACCAAGAGTATGCTTACAGGGCCAGAGTTGACATGAATTT	927
Qy	870	CATTGGAAGAGATACCAAGCGCAGGAACAGCATTCCTTTGAGAAAGCTATTTACCAAGA	929
Db	928	TATATAGTAGGAGGTACCAAGAGGAGAAATAGCATCCCTCGCATCGTGCATTTGCAAGA	987
Qy	930	CACCTGTTGAGATTACGAGAAAGTCTGTCGCACTTCTCGGTGAAGATGATGCTATTC	989
Db	988	CACCTACTGAGACTATATGAAAAATGCTTGGCATTTGATGGCCAGTCGAGGCTTGATT	1047
Qy	990	AATCAATCTTC 1000	
Db	1048	TACAAGTACTC 1058	
RESULT 13			
ADB94943			
ID ADB94943 standard; cDNA; 1293 BP.			
XX ADB94943;			
XX 04-DEC-2003 (First entry)			
XX Programmed cell death pathway protein annexin cDNA #1.			
XX			
KM programmed cell death; plant development; plant cell cycle; AT12; DAD1;			
KM Dnaee; 11s; 1sd1; nucellin-like aspartic protease; annexin; prohibitin;			
KM fen-like protein; rsc2; retinoblastoma-related protein; SIN3; TFIID;			

KM	TEGT; xylogenic Riase; pur-alpha; cyteine protease; RPPs-like protein;
KM	gp 91 NADPH oxidase subunit; NRP-like protein; BaG-1;
KM	defender against cell death; lethal leaf spot; lesion stimulating death;
KM	seven in absentia; transcription initiation factor;
KM	teatle enhanced gene transcript; gene; ss.
OS	Pinus radiata.
XX	US2003082724-A1.
PN	01-MAY-2003.
PD	14-AUG-2002; 2002US-00219220.
PF	04-JUN-1999; 99US-00325932.
XX	(GENE-) GENESIS RES & DEV CORP LTD.
XX	Pinm B, Laeham A;
XX	WPI; 2003-786916/74.
DR	New isolated polynucleotide useful for modulating programmed cell death,
XX	altering the development cycle of plant cells, and subsequently modifying
PT	plant development.
PS	Claim 1; Page 28-29; 214pp; English.
XX	The invention describes an isolated polynucleotide (I) comprising a
CC	nucleotide sequence that is one of 145 fully defined sequences of 221-
CC	3415 base pairs (bp), given in the specification, and/or its complements,
CC	reverse complements, reverse sequences, or sequences having 75, 90 or 95
CC	% sequence identity to or that hybridise under stringent hybridisation
CC	conditions to one of the 145 sequences. The methods and compositions of
CC	the present invention to do with (I) are useful for modulating programmed
CC	cell death and thereby altering the development cycle of plant cells, and
CC	altering plant development. This sequence encodes a protein associated
CC	with the programmed cell death pathway.
XX	Sequence 1293 BP; 359 A; 290 C; 313 G; 331 T; 0 U; 0 Other;
SO	
	Query Match 38.1%; Score 468.6; DB 10; Length 1293;
	Best Local Similarity 68.4%; Pred. No. 2.5e-109;
	Matches 664; Conservative 0; Mismatches 304; Indels 3; Gaps 1
OY	30 AAAATGGCGACTCTTAAAGTTTCTGAATCTGTTCTGCTCTTGAATGCTGAGCA 89
DB	91 AACAAATGCGACTCTCAACCGTCCCGAGCACTGCCCTGTAGCCGATGACTGAGCA 150
OY	90 ATTGAGAACCGCTTTTGAAGATGGGGTACGAGAGCACTTGAATCATATCATTTGGC 149
DB	151 GCTCCGAGACGCTTTCGAGATGGGAAACAAATGAGAGCTGATCATATTCATATTTGG 210
OY	150 TCACAGAGTGTCTGAACAGAGAAAGTATCAGGCAAGCATATACAGCAAAACTACGGCGA 209
DB	211 TCATAGAGATGCGCGCGCAGAGAGACTGATTCGGCAACTATGCCAGACTTACGGCGA 270
OY	210 AGACTTTCTCAAGACTCTTGACAGAGAGCTCTTAACGATTGTGAGAGAGCTATCTTGT 269
DB	271 GGAAGCTCTTCAAGGCAATTTGACAGAACTTAACCAATGATTTTCAGAGGCTGTGGTCT 310
OY	270 GTGGAATCTTGAACCCGGTGAACCGTGAATCTTTATTTGGCTAATGAACCTTCAAAAAGAT 329
DB	331 TTGATCACTTGATTCGGCTGAACGTGAATCGTATGCGATATACGACGAAAGATG 390
OY	330 GACTTCAAGCAACCAAGTTCTTATGGAAGTTGCTTTCACAGACATCAACGAGCTGCT 389
DB	391 GACTTCAAGCAACGAGTTCTCATATGAAATAGCTGACAGAGCTCTCCGACAGACTTGT 450
OY	390 TCAAGCTAGGCAAGCTTACATGATCTGCTACAGAAAGTCTTCTTGAAGAGAGCTTGCTCA 449
DB	451 TATGCGAAGCAAGCAAGTATATCATGCCGATTAACAGAAAGTCAATGAGAGAGCTGCTCA 510



PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147382P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155466P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 08-OCT-1999; 99US-0158029P.  
PR 12-OCT-1999; 99US-0158232P.  
PR 13-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 23-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161929P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 37.2%; Score 457.2; DB 3; Length 1160;  
Best Local Similarity 67.3%; Pred. No. 2e-106;  
Matches 645; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 30 AAAATGGGACCTTAAGGTTCTGATTCGTGTCGCTTCCTGATGATGCTGACA 89  
Db 92 AACATGGGCTCTCTAAAGTCCCAAGCAATGTTCTTCCGAAGATGACCCGAGCA 151  
QY 90 ATTGAGACCGCTTTTGAAGATGGGTACGAAGAGACTTATCATATCAATCTTGGC 149  
Db 152 ACTCCAAAGGCTTTTTCAGATGGGTACCAAGAGAGCTATCATATCAATCTTACG 211  
QY 150 TCACAGATGCTGAACAGAGAAATCATCAGGCAAGCATACCAAGAACTTACGGCA 209  
Db 212 TCACAGAAAGCAGCAACGCGCTTATCCGACGTTTAAAGCATTACATCAATGA 271  
QY 210 AGACCTTTCAGACCTCTTGAAGAGCTCTCTAAGATTTGAGAGACTATCTTGT 269  
Db 272 GGAATCTTTCAGACCTTGAAGAGCTTTCAGACCTTTCAGAGAGCTGTGATGTT 331  
QY 270 GTGACTCTTGAACCGGATGAGCGTATGCTTATTTGGCTAATGAAGTACAAAAGATG 329  
Db 332 GTGACTCTTGAATCCAGAGAGAGATGCTTATTTGGCTAAGATTCACCAAGATGT 391  
QY 330 GACTTCAAGCAACCAAGTCTTATGGAAGTGTGCTGACAGGACATCAACGAGCTGCT 389  
Db 392 CACCAAGAACATGAGGTTCTTGTGAAATCGCTTCAAGGCTGCTTGAAGCTTAT 451  
QY 390 TCACGCTAGGCAAGCTTACATGCTGCTCAAGAAAGTCTTGAAGAGAGCTTGTCTCA 449  
Db 452 CAAGTCAAGCAAGCTTACAGCTGATACAAAGAAATCAATGAGAGAGATGTCGCA 511  
QY 450 CCACACTACCGGACTTCAAGAAAGCTTTGGTTCTCTTGAACCTCATACAGTATCA 509  
Db 512 AACACATCTGGGAGCTTCTGTAAGCTTGTGCTCTCTTGAAGACTTCAAGATATGA 571  
QY 510 AGGAGTGAAGTGAACATGACATGCTGCTAAGCAAGAGCTTAAGCTGCTCAATGAGAAAT 569  
Db 572 AGGAGTGAAGTGAACATGATGCTTGAAGATCTGAAGCTTAAGATCTTCAAGAGAGGT 631  
QY 570 CAAGGACAGCACTAACATGATGAGGATGTTATTTGAATCTTGCACAAAGAGCAAGC 629  
Db 632 CTCAGAGAAATCTTACATGACATGACATCTTCAATCGAATCTTGAACAAAGAGCAAGC 691  
QY 630 TCAGATCAATGCTACTTTTAACTGTTACCGTAAAGATGATGCGAGGAATTTCTCAAGG 689  
Db 692 ACAGCTCGGTGCAACCTACACCACTACCAACAAAGATGATGAAAGCGCATTAAACAAGA 751

QY 690 TCTTGAGAGAGATGATGATGACAAAGTCTTTCGACCTTTGAGTCAACATTCAGTG 749  
Db 752 CTTGAGAGAGAGATCGGACACAATGACTACATGAAACTGTAAGCTTAATCATG 811  
QY 750 CTTGACAAAGCCAGAGCTTTACTTTGTGATGTTCTTTCGTTCCGACATCAACAAAACCTGG 809  
Db 812 TTGACATACCTCGAGAGCATTTTGAGAGGTTCTTCGTCATCAATCAACAAATGG 871  
QY 810 AACTGATGAGAGAGCATCTGATGATGATGACCAAGAGCTGATGATGACTTGAAGGT 869  
Db 872 AACAGAGATGGGGCTAACCCGAGTCGACTTACAGAACTGAAAGTTGACATGGAACG 931  
QY 870 CATTGAGAGAGATGACAGCGAGAAACAGCATTTCTTTCGAGAAAGCTATTACCAAGA 929  
Db 932 CATCAAGAGATATCAGCGAAAGAACGACATTTCTTTCGACCGTCCTATCGCCAAAGA 991  
QY 930 CACTCGTGAAGATTACGAGAAAGTCTCGTGACATTTTCGTTGAAGATGATGCTTAA 987  
Db 992 CACTTCGTGACTATGAGACATGCTTGTGCTTTCGTCGACATGCGATGCTTGA 1049

## RESULT 15

AAC49728

ID AAC49728 standard; DNA; 1156 BP.

XX AAC49728;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 62219.

XX Hybridisation assay; genetic mapping; gene expression control;  
KM protein identification; signal transduction pathway; metabolic pathway;  
KM promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0127462P.

PR 01-APR-1999; 99US-0128234P.

PR 06-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132407P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134321P.

PR 18-MAY-1999; 99US-0134370P.

PR 19-MAY-1999; 99US-0134768P.

PR 20-MAY-1999; 99US-0134941P.

PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 14-JUN-1999; 99US-0138847P.  
PR 16-JUN-1999; 99US-0139119P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 19-JUL-1999; 99US-0144336P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145090P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145226P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148655P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159283P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.

PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 37.0%; Score 455.6; DB 3; Length 1156;  
Best Local Similarity 67.2%; Pred. No. 5.1e-106;  
Matches 644; Conservative 0; Mismatches 314; Indels 0; Gaps 0;

QY 30 AAAAAAGCGACTCTTAAGGTTCTGATTCGTGCTCTTCTGATGATGCTGACGA 89  
DB 91 AACATGGGCTCTCTCAAGTCCAGCAAGCATGTTCTCTCCGGAAGATGACGCCAGCA 150  
QY 90 ATTGAGAACCGCTTTTGAAGATGGGGTACGAACGAGACTTATCATATCAATCTTGGC 149  
DB 151 ACTCCAAAGGCTTTTTCAGAGTGGGTACCAACGAAGCTGATCATATCAATCTACGC 210  
QY 150 TCACAGAAAGTGTGAACAGAGGAAGTCATCAGGCAAGCATCCAGAAACCTAAGCGCA 209  
DB 211 TCACAGAAAGCGACGAACAGCGACTTGATCCGACGCTTTATGCACTTACATGA 270  
QY 210 AGACTTCTCAAGACTCTTGACAAAGGCTCTTAACGATTTGAGAGCTATCTTGT 269  
DB 271 GATCTTCTCAAGCATTTAGACAAAGCTTCTTACGACTTTGAGAGCTGTGATGT 330  
QY 270 GTGGACTCTTGAACCCGGTGAAGCTGATGCTTATTTGCTAATGAAGCTAAGAAAGATG 329  
DB 331 GTGGACTCTTGAATCCACAGAGAGATGCTTATTTGGCTAAGATCCACCAAGATGT 390  
QY 330 GACTTCAAGCAACCAAGTCTTATGGAAGTGTGCAAGAGCATCAACGAGCTGCT 389  
DB 391 CACCAAGAAACAAATGGGTTCTGTGAATCGTTGCAAGAGCTGCTTGAAGCTTAT 450  
QY 390 TCACGCTAGCAAGCTTACCATGCTCGTACAAAGATCTTGAAGAGAGCTTCTCA 449  
DB 451 CAAGTCAAGCAAGCTTACCAAGCTCGATACAAAGAAATCAATGAGAGATGTCGCGCA 510  
QY 450 CCACACTACCGGTGACTTCAAGAAAGTTTGGTTCTCTTGTACTCATACAGGTACGA 509  
DB 511 ACACACATCTGGGACCTTCGTAAAGCTCTGCTTCTTGTGAGCACTTCCAGGTATGA 570  
QY 510 AGGAGATGAAGGAACATGACATTTGGCTAAGCAAGAAAGCTAAGCTGCATAGAAAT 569  
DB 571 AGGAGATGAATGGAATATATGCTTGCAAGATCTGAAGCTAAGATCTTCAAGAAAGGT 630  
QY 570 CAAGCAAGCACTAACATGATGAGATGTTATGAATCTTGCACAAAGCAAGC 629  
DB 631 CTCAGAGAAATCTTACAGTGAAGATGACTTCAAGAACTTTGACAAAGAAAGCAAGC 690  
QY 630 TCAGATCAATGCTACTTTTAACCGTTACCAAGATGATGAGCGAGAAATTCAGAG 689  
DB 691 ACAGCTCGGTGCAACCTAACCTACCAACAGATGGAAGAAAGCATTAACAGAA 750  
QY 690 TCTTGAAGAAAGAGATGATGACAAAGTTCCTTGAACCTTTGAAGTCAACATTCAGT 749  
DB 751 CTTGAAGAAAGTGGAGCAACATGATCAATGAATCACTAAGAGCTGATATCAATG 810  
QY 750 CTTGACAAAGACGAGCTTACTTTGTCATGTTCTTCCGTTGACGATCAACAAATCG 809  
DB 811 TTTGACATACCTTGAGAGACATTTTGAAGAGTTCTTCCTCTCATATCAACAAATGG 870  
QY 810 AACGTGAAGAGACATCTAGATTTGTGACCAACAGAGCTGATGACTTGAAGT 869  
DB 871 AACAGAGATGGGAGCTTAACCGAGTGTGACTACACAACTGAAGTTGACATGGAAG 930  
QY 870 CATTGAGAGAGTACCGCGGAGAGACAGATTCCTTTGGAAGAAAGCTATTAACAAAGA 929  
DB 931 CATCAAGAGAAATACAGCGAAAGAAACAGATTCCTTTGACCGGTGATGCAAGAA 990  
QY 930 CACTGTGAGATTACGAGAAAGATGCTGTCACCTTCTCGGTGAAGATGATGTTAA 987  
DB 991 CACTTGTGTACTATGAGAGACATGCTTGTGCTTCTTCGGAACATGGGATGCTTGA 1048



Tue Aug 23 09:03:46 2005

us-10-690-564-1.rng

Page 20

Search completed: August 22, 2005, 07:25:19  
Job time : 813 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 04:40:07 ; Search time 4649 Seconds

(without alignments)  
10070.779 Million cell updates/sec

Title: US-10-690-564-1

Perfect score: 1230  
Sequence: 1 cccgcgcgcgcgaacactaa.....aaaaaaaaaaaaaaaaaaaaa 1230

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1108.6	90.1	1155	3	CNSOAB51
2	1081.4	87.9	1095	3	CNSOABR2
3	1058	86.0	1137	3	CNSOABTU
4	1030	83.7	1063	5	EX838972
5	947.6	77.0	966	3	CNSOABV3
6	924.2	75.1	1034	3	CNSOABY3M
7	892.2	72.5	910	7	CF652924
8	836.2	68.0	865	3	CNSOABCI1
9	831.8	67.6	843	7	CF652048
10	786.2	63.9	814	6	CA782049
11	743.6	60.5	782	5	BU635010
12	734.4	59.7	775	5	BU636350
13	703	57.2	703	1	AV784024
14	700	56.9	702	1	AV785121
15	667.8	54.3	679	5	EX836971
16	652.4	53.0	670	3	CNSOABT2
17	649.4	52.8	672	5	BP561140
18	641.4	52.1	660	6	CB256182
19	630.4	51.3	676	2	BE037587
20	627.2	51.0	661	5	BP560951
21	608.4	48.9	665	5	BP560725
22	601.4	49.5	627	1	AV782813
23	594	48.3	594	1	AV559022
24	590.4	48.0	592	1	AV538505

C	25	568.4	46.2	583	1	AI998553	AI998553
	26	543.4	44.2	578	6	CB256316	CB256316
	27	540.6	44.0	576	7	CF652799	CF652799
	28	540.2	43.9	569	6	CB259325	CB259325
	29	538.6	43.8	720	6	CD815439	CD815439
	30	535	43.5	535	9	CNSO058J	AL088081
	31	512.2	41.6	667	6	CD816318	CD816318
	32	505	41.1	676	6	CD818668	CD818668
	33	504.2	41.0	673	6	CD829166	CD829166
	34	502.4	40.8	518	6	CB257482	CB257482
	35	500.4	40.7	767	6	CD814071	CD814071
	36	500	40.7	500	1	AV550330	AV550330
	37	499.4	40.6	515	6	CB257484	CB257484
C	38	489	39.8	489	1	AV544263	AV544263
	39	488.8	39.7	426	6	CB256030	CB256030
	40	485.8	39.5	929	6	CB350421	CB350421
	41	475.2	38.6	640	6	CD820443	CD820443
	42	469	38.1	889	7	CO121757	CO121757
C	43	463	37.6	463	5	BP572264	BP572264
C	44	463	37.6	464	5	BP581497	BP581497
C	45	462.2	37.6	468	5	BP570388	BP570388

#### ALIGNMENTS

RESULT 1  
LOCUS  
CNSOAB51  
DEFINITION  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GS17S16ZB06 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION  
BX818120  
VERSION  
BX818120.1  
KEYWORDS  
HTC; GSUT cDNA.  
SOURCE  
Arabidopsis thaliana  
ORGANISM  
Arabidopsis thaliana

REFERENCE  
AUTHORS  
Casati, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schacher, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.,  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

#### TITLE

JOURNAL  
REFERENCE  
Unpublished  
2 (bases 1 to 1155)

#### AUTHORS

Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

#### COMMENT

The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Casati V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schacher V., Weissenbach J., Salanoubat M.  
URG RNA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences) . 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggbsource=Arabidopsis.

#### FEATURES

source  
1..1155  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"

gene  
/clone="GSLT51L62ZB06"  
/issue\_type="Silique"  
/plasmid="pCMVSPORT\_6"  
1..1155  
/gene="At1g35720"

## ORIGIN

Query Match 90.1%; Score 1108.6; DB 3; Length 1155;  
Best Local Similarity 98.6%; Pred. No. 1.1e-278;  
Matches 1139; Conservative 0; Mismatches 14; Indels 2; Gaps 2;

QY 21 AAGTAGAAGAAAATGGCGACTCTTAAGGTTTCTGATTCCTGCTCTCTTCTGATGA 80  
DB 1 AAGTAGAAGAAAATGGCGACTCTTAAGGTTTCTGATTCCTGCTCTCTTCTGATGA 60  
QY 81 TGTGTAGACATTGAGAACCGCTTTTGAAGATGGGGTACGAAAGAGACTTATCTATTC 140  
DB 61 TGTGTAGACATTGAGAACCGCTTTTGAAGATGGGGTACGAAAGAGACTTATCTATTC 120  
QY 141 AATCTTGGCTCACAAGAGTGTGAACAGAGAAAGTCAATCAGAGCAATACCAAGAAC 200  
DB 121 AATCTTGGCTCACAAGAGTGTGAACAGAGAAAGTCAATCAGAGCAATACCAAGAAC 180  
QY 201 CTACGGCGAAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTACGATTTGAGAGACC 260  
DB 181 CTACGGCGAAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTACGATTTGAGAGAAC 240  
QY 261 TATCTTGGTGTGACCTTGAACCCGGTGAAGCTGATGCTTTATGGCTTAATGAACCTAC 320  
DB 241 TATCTTGGTGTGACCTTGAACCCGGTGAAGCTTATGCTTTATGGCTTAATGAACCTAC 300  
QY 321 AAAAAGATGACTTCAAGCAACCAAGTCTTATGAAGTTGCTTGAACAAGACATCAAC 380  
DB 301 AAAAAGATGACTTCAAGCAACCAAGTCTTATGAAGTTGCTTGAACAAGACATCAAC 360  
QY 381 GCAGCTGCTTCAAGCTGAGCACTTACATGCTCGCTACAGAAAGTCTTGAAGAGGA 440  
DB 361 GCAGCTGCTTCAAGCTGAGCACTTACATGCTCGCTACAGAAAGTCTTGAAGAGGA 420  
QY 441 CGTGTGCTACCACTACCGGTGACTTTCGAAGGCTTTTGTCTTGTATACCTCATA 500  
DB 421 CGTGTGCTACCACTACCGGTGACTTTCGAAGGCTTTTGTCTTGTATACCTCATA 480  
QY 501 CAGGTACGAAGAGATGAAGTGAACATGATGCTTAAGCAAGAGTAAAGCTGTGCA 560  
DB 481 CAGGTACGAAGAGATGAAGTGAACATGATGCTTAAGCAAGAGTAAAGCTGTGCA 540  
QY 561 TGAAGAAATCAGAGACAGACATCAATGATGAGATGTTATTAGATCTTGTCCAAAG 620  
DB 541 TGAAGAAATCAGAGACAGACATCAATGATGAGATGTTATTAGATCTTGTCCAAAG 600  
QY 621 AAGCAAGCTCAGATCAATGCTACTTTTAAACGTTTCAAGATGATCAAGGAGGAAT 680  
DB 601 AAGCAAGCTCAGATCAATGCTACTTTTAAACGTTTCAAGATGATCAAGGAGGAAT 660  
QY 681 TCTCAAGAGCTTGAAGAGAGATGATGACAAGTCTTGAACCTTTTGAAGTCAAC 740  
DB 661 TCTCAAGAGCTTGAAGAGAGATGATGACAAGTCTTGAACCTTTTGAAGTCAAC 720  
QY 741 CATTCACTGCTTGAACAAGACAGAGCTTTTCTTGTGATGTTCTTGTTCAGCAATCA 800  
DB 721 CATTCACTGCTTGAACAAGACAGAGCTTTTCTTGTGATGTTCTTGTTCAGCAATCA 780  
QY 801 CAAAACCTGGAACGATGAAGAGAGACACTCACTAGAAATTTGACCAAGAAGCTGAGATTGA 860  
DB 781 CAAAACCTGGAACGATGAAGAGAGACACTCACTAGAAATTTGACCAAGAAGCTGAGATTGA 840  
QY 861 CTGGAAGCT-CATTGAGAGAGATACCAAGCGCAGAAACAGCATTCCTTTGAGAAAGCTA 919  
DB 841 CTGGAAGCTCATTGAGAGAGATACCAAGCGCAGAAACAGCATTCCTTTGAGAAAGCTA 900  
QY 920 TTACCAAGACACTCTGTGAGATTACGAGAAAGATGCTGTGCACTTCTCGTGAAGATG 979

DB 901 TTACCAAAAAAACTGTGAGATTACGAGAAAATGCTCTCGCACTTCTCGTGAAGATG 960  
QY 980 AGCTTAATCATCATTCCTCCACAGAGAAACATAGCGCTCTACAGCTTCTGTTATCT 1039  
DB 961 AGCTTAATCATCATTCCTCCACAGAGAAACATAGCGCTCTACAGCTTCTGTTATCT 1020  
QY 1040 CTTATCTCCCTCTCTCTCTCTTGTGATGATTCGAATGCTTGAATTTGTTTCTACAAA 1099  
DB 1021 CTTTCTCCCTCTCTCTCTCTTGTGATGATTCGAATGCTTGAATTTGTTTCTACAAA 1080  
QY 1100 ACCCTGTTGTTCTGTTGTGTTTGAAGTCTTAA-ATAATGCAAAAGAGAGACAG 1158  
DB 1081 ACCCTGTTGTTTGTGTTGTGTTGATGTTCTTAATATATGCAAAAAGAGAGACAG 1140  
QY 1159 AGAGAACCAAGTGTG 1173  
DB 1141 AGAGAACCAAGTGTG 1155

RESULT 2  
CNS0ABR2  
LOCUS  
DEFINITION  
CNS0ABR2 1095 bp mRNA linear HTC 06-PEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLT51L62G07 of Hormone Treated Callus of strain col-0 of  
Arabidopsis thaliana (thale cress).  
BX816579.1 GI:42471684

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
HTC; GSLT; cDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 1095)

REFERENCE  
AUTHORS  
TITLE  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Unpublished  
2 (bases 1 to 1095)  
Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr)

COMMENT  
- Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences) . 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
length

http://www.genoscope.cns.fr/cgi-bin/g9b/g9b?source=Arabidopsis.  
location/Qualifiers  
1..1095

## FEATURES

source

/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLT51L62G07"  
/issue\_type="Hormone Treated Callus"  
/plasmid="pCMVSPORT\_6"  
1..1095  
/gene="At1g35720"

## ORIGIN

Query Match 87.9%; Score 1081.4; DB 3; Length 1095;  
 Best Local Similarity 99.8%; Pred. No. 1.4e-271;  
 Matches 1093; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 12 AAACACTAAAAAGTGAAGAAAAATGGGACTCTTAAGGTTCTGATTTCTGTCTGCC 71  
 DB 1 AAACACTAAAAAGTGAAGAAAAATGGGACTCTTAAGGTTCTGATTTCTGTCTGCC 60

QY 72 TTCGATGATGCTGACGAATTGAGAACCGCTTTTGAAGATGGGGTACGAACGAGACT 131  
 DB 61 TTCGATGATGCTGACGAATTGAGAACCGCTTTTGAAGATGGGGTACGAACGAGACT 120

QY 132 GATCATATCAATCTTGGCTCAGAGAAGTGTGAACAGAGGAAAGTCAATCAGGACGATA 191  
 DB 121 GATCATATCAATCTTGGCTCAGAGAAGTGTGAACAGAGGAAAGTCAATCAGGACGATA 180

QY 192 CCAAGAAACCTACGGCGAAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTACGATTT 251  
 DB 181 CCAAGAAACCTACGGCGAAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTACGATTT 240

QY 252 CGAGAGGCTATCTTGTGTGAGACTCTTGAACCGGTGAGCGTGAATCTTATGGCTAA 311  
 DB 241 CGAGAGGCTATCTTGTGTGAGACTCTTGAACCGGTGAGCGTGAATCTTATGGCTAA 300

QY 312 TGAAGCTACAAAAAGATGAGACTTCAAGCAACCAAGTCTTATGGAAGTTGCTGACAA 371  
 DB 301 TGAAGCTACAAAAAGATGAGACTTCAAGCAACCAAGTCTTATGGAAGTTGCTGACAA 360

QY 372 GACATCAACGACGCTGCTTCAACGCTAGCAAGCTTACCATGCTGCTACAAAGAGTCT 431  
 DB 361 GACATCAACGACGCTGCTTCAACGCTAGCAAGCTTACCATGCTGCTACAAAGAGTCT 420

QY 432 TGAAGAGAGCTTGTCTCAACCACTACCGGTGACTTCAAGAAAGTTTGGTTCTCTGT 491  
 DB 421 TGAAGAGAGCTTGTCTCAACCACTACCGGTGACTTCAAGAAAGTTTGGTTCTCTGT 480

QY 492 TACCTCATACAGGTGAGAGAGATGAGAGATGAGAAATGGCTAAGCAAGAGTAA 551  
 DB 481 TACCTCATACAGGTGAGAGAGATGAGAGATGAGAAATGGCTAAGCAAGAGTAA 540

QY 552 GCTGTGCTCATGAGAAAAATCAAGAGCAAGCACTACATATGATGAGATTTATTAAGATCT 611  
 DB 541 GCTGTGCTCATGAGAAAAATCAAGAGCAAGCACTACATATGATGAGATTTATTAAGATCT 600

QY 612 GTCCACAAGAGCAAGAGCTGATGATCAATGCTACTTTTAAACCGTTTACCAAGATGATC 671  
 DB 601 GTCCACAAGAGCAAGAGCTGATGATCAATGCTACTTTTAAACCGTTTACCAAGATGATC 660

QY 672 CGAGAAATCTCAAGAGCTTGAAGAGAGATGAGAGATGAGAAAGTTCTTGTGACTTTT 731  
 DB 661 CGAGAAATCTCAAGAGCTTGAAGAGAGATGAGAGATGAGAAAGTTCTTGTGACTTTT 720

QY 732 GAGGTCAACCATTCAGTGTGACAAAGACAGAGCTTACTTTGTGATGTTCTTGTGCTTC 791  
 DB 721 GAGGTCAACCATTCAGTGTGACAAAGACAGAGCTTACTTTGTGATGTTCTTGTGCTTC 780

QY 792 AGCAATCAACAAAACTGGAACCTGATGAGAGAGCACTCACTAGAAATTTGACCAAGAGC 851  
 DB 781 AGCAATCAACAAAACTGGAACCTGATGAGAGAGCACTCACTAGAAATTTGACCAAGAGC 840

QY 852 TGAAGATTGACTTGAAGGTCAATGAGAGAGTACAGCGGAGAGAAACGACTTCTTGTGA 911  
 DB 841 TGAAGATTGACTTGAAGGTCAATGAGAGAGTACAGCGGAGAGAAACGACTTCTTGTGA 900

QY 912 GAAAGCTATTAACAAAGACACTCGTGAAGATTACGAAGAAATGCTCTGCCACTTCTCG 971  
 DB 901 GAAAGCTATTAACAAAGACACTCGTGAAGATTACGAAGAAATGCTCTGCCACTTCTCG 960

QY 972 TGAAGATGATGCTTATCAATCAATCTTCAACAGAGAAACATTAAGCTGCTTACAGTTC 1031  
 DB 961 TGAAGATGATGCTTATCAATCAATCTTCAACAGAGAAACATTAAGCTGCTTACAGTTC 1020

QY 1032 TGTTATCTCTATCTCCCTCTCTCT-CTTTGATGAGTTTCAAAATGTTTGAATTTGTT 1090

DB 1021 TGTTATCTCTATCTCCCTATCTCTCTGCTTGTATGATGTTCAAAATGTTGATTTGTT 1080

QY 1091 TCTACAAAAACCTTG 1105

DB 1081 TCTACAAAAACCTTG 1095

RESULT 3  
 CDS0ABTU 1137 bp mRNA linear HTC 06-FEB-2004  
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 DEFINITION GSUTRGH182F02 of Hormone Treated Callus of strain col-0 of  
 Arabidopsis thaliana (thale cress).  
 BX815936  
 ACCESSION BX815936.1 GI:42472184  
 VERSION HTC; GSUT cDNA.  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 1137)  
 Castelli,V., Aury,J.M., Jallion,O., Winkler,P., Scarpelli,C., Schachter,V.,  
 Menard,M., Cruaud,C., Quetier,F., Salanoubat,M.,  
 Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.,  
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome  
 Annotation  
 Unpublished  
 2 (bases 1 to 1137)  
 JOURNAL Genoscope.  
 REFERENCE Direct Submission  
 AUTHORS Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 JOURNAL - Web : www.genoscope.cns.fr)  
 COMMENT The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli  
 V., Aury J.M., Jallion O., Winkler P., Menard M., Cruaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information Center for Protein  
 Sequences) . 5 prime and 3 prime are assembled with phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
 length  
 http://www.genoscope.cns.fr/cgi-bin/gsb/gsb?source=Arabidopsis.  
 FEATURES  
 source  
 1..1137  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSUTRGH182F02"  
 /issue\_type="Hormone Treated Callus"  
 /plasmid="pCMVSPORT\_6"  
 1..1137  
 /gene="Atlg35720"

ORIGIN  
 Query Match 86.0%; Score 1058; DB 3; Length 1137;  
 Best Local Similarity 97.8%; Pred. No. 1.9e-265;  
 Matches 1105; Conservative 0; Mismatches 20; Indels 5; Gaps 3;

QY 12 AAACACTAAAAAGTGAAGAAAAATGG-CGACTCTTAAGCTT-TGTGATTTCTGTCTCG 67  
 DB 2 AAACACTAAAAAGTGAAGAAAAATGAAGACTCTTTAAGAAATAAAAAATCTGTTCCG 61

QY 68 CTCCTCTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGGTACGAAGAGAG 127  
 DB 62 CTCCTCTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGGTACGAAGAGAG 121

```
QY 128 ACTGATCATATCATCTTGGCTCAGAGAACTGCTGAACAGAGAAAGTCATCAGGCGAAG 187
    |||
Db 122 ACTGATCATATCATCTTGGCTCAGAGAACTGCTGAACAGAGAAAGTCATCAGGCGAAG 181
QY 188 CATACACAGAAACCTACGCGGAAGACCTTCTCAAGA CTTTGAACAAGAGCTCTCTAAGC 247
    |||
Db 182 CATACACAGAAACCTACGCGGAAGACCTTCTCAAGA CTTTGAACAAGAGCTCTCTAAGC 241
QY 248 ATTTCAGAGAGCTATCTTGTGTGAAGCTTGAACCGGTGAGCGGTGAGCTTATTTGG 307
    |||
Db 242 ATTTCAGAGAGCTATCTTGTGTGAAGCTTGAACCGGTGAGCGGTGAGCTTATTTGG 301
QY 308 CTATAGAGCTACAAAAGATGAGCTTCAAGCAACCAAGTCTTATGAAAGTTGCTTGA 367
    |||
Db 302 CTATAGAGCTACAAAAGATGAGCTTCAAGCAACCAAGTCTTATGAAAGTTGCTTGA 361
QY 368 CAAGAATCAACGACGCTGCTTACGCTAGGCAAGCTTACATGCTGCTACAGAAGT 427
    |||
Db 362 CAAGAATCAACGACGCTGCTTACGCTAGGCAAGCTTACATGCTGCTACAGAAGT 421
QY 428 CTCTTAAAGAGAGCTTGTGCTACACACATACCGGTGACTTCAAGAGCTTTGGTTTC 487
    |||
Db 422 CTCTTAAAGAGAGCTTGTGCTACACACATACCGGTGACTTCAAGAGCTTTGGTTTC 481
QY 488 TTGTTACTCATACAGGTACGAAGAGATGAATGAAACATGACATTGGCTAAGCAGAAG 547
    |||
Db 482 TTGTTACTCATACAGGTACGAAGAGATGAATGAAACATGACATTGGCTAAGCAGAAG 541
QY 548 CTAAAGCTGTGCATGAGAAATCAAGACATCAATGATGAGATGTTATTAGAA 607
    |||
Db 542 CTAAAGCTGTGCATGAGAAATCAAGACATCAATGATGAGATGTTATTAGAA 601
QY 608 TCTGTGCCAAGAAGAAAGCTCAGATCAATGCTTAAACGTTTACCAAGATGATC 667
    |||
Db 602 TCTGTGCCAAGAAGAAAGCTCAGATCAATGCTTAAACGTTTACCAAGATGATC 661
QY 668 ATGGCAGAGAAATTTCTCAAGAGCTTGAAGAGAGATGATGACAAGTTCTTGCAC 727
    |||
Db 662 ATGGCAGAGAAATTTCTCAAGAGCTTGAAGAGAGATGATGACAAGTTCTTGCAC 721
QY 728 TTTTGAAGTCAACCAATTCAGTCTTGAACAAGACCAAGCTTTATCTTGTGATGTTCTTC 787
    |||
Db 722 TTTTGAAGTCAACCAATTCAGTCTTGAACAAGACCAAGCTTTATCTTGTGATGTTCTTC 781
QY 788 GTTCACCAATCAACAAACCTGGAACCTGATGAAGAGACCTCAATGATTTGACCAAA 847
    |||
Db 782 GTTCACCAATCAACAAACCTGGAACCTGATGAAGAGACCTCAATGATTTGACCAAA 841
QY 848 GAGCTGAGATTGACTTGAAGGTCAATGAGAGAGTACACGCGCAAGAACAGATTCCTT 907
    |||
Db 842 GAGCTGAGATTGACTTGAAGGTCAATGAGAGAGTACACGCGCAAGAACAGATTCCTT 901
QY 908 TGAAGAAAGCTATTAACAAAGACATCTGTGAGATTAACGAAGATGCTGTGCGACTTC 967
    |||
Db 902 TGAAGAAAGCTATTAACAAAGACATCTGTGAGATTAACGAAGATGCTGTGCGACTTC 961
QY 968 TCGGTAAAGATGATGCTTAATCAATCAATCTCCACAGAGAAACATTAAGTGTCTACAG 1027
    |||
Db 962 TCGGTAAAGATGATGCTTAATCAATCAATCTCCACAGAGAAACATTAAGTGTCTACAG 1021
QY 1028 CTCTGTATATCTTATATCTCCCTCTCTCTCTCTTGTGAGAGTTCAAAATGTTGATTT 1087
    |||
Db 1022 CTCTGTATATCTTATATCACTCTGTCTATCTATGATGAGTTTCAAAATGTTGATTT 1081
QY 1088 GTTTCACAA-AAAACCTGTGTTGTGTTGTGTTGTGTTGAGTCTTAA 1136
    |||
Db 1082 TTTTATATATAAACCTGTGTTGTGTTGTGTTGTGTTGAGTCTTAA 1131
RESULT 4
BX838972 1063 bp mRNA linear EST 11-FEB-2004
```

```
DEFINITION BX838972 Arabidopsis thaliana Flowers and buds Col-0 Arabidopsis
thaliana cDNA clone GSTLFB502D03 5PRIM, mRNA sequence.
ACCESSION BX838972
VERSION BX838972.1 GI:42533055
KEYWORDS EST
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1063)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Queller,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished (2004)
JOURNAL Contact: Genoscope
COMMENT Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences).
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST
http://www.genoscope.cns.fr/cgi-bin/ggb/ggbsource-Arabidopsis.
location/Qualifiers
source 1..1063
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotYPE="Col-0"
/db_xref="taxon:3702"
/clone="GSTLFB502D03"
/cisue_type="Flowers and buds"
/clone_lib="Arabidopsis thaliana Flowers and buds Col-0"
ORIGIN
Query Match 83.7%; Score 1030; DB 5; Length 1063;
Best Local Similarity 99.0%; Pred. No. 3.9e-258;
Matches 1036; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 108 AGATGGGGTACGAACGAGAGCTTATCATATCAATCTTGGCTCAGAGAGTGTGAACA 167
    |||
Db 1 AGATGGGGTACGAACGAGAGCTTATCATATCAATCTTGGCTCAGAGAGTGTGAACA 60
QY 168 GAGGAAAGTCATCAGGCAAGCATATCACGAAACCTACGCGCAAGACCTTCTCAAGCTCT 227
    |||
Db 61 GAGGAAAGTCATCAGGCAAGCATATCACGAAACCTACGCGCAAGACCTTCTCAAGCTCT 120
QY 228 TGAACAAGAGCTCTCTAAGATTTTGAAGAGACTATCTTGTGTGAGACTCTGAACCCGG 287
    |||
Db 121 TGAACAAGAGCTCTCTAAGATTTTGAAGAGACTATCTTGTGTGAGACTCTGAACCCGG 180
QY 288 TGAGCGTATGCTTATATTTGGCTAATGAGCTCAAAAAGATGAGACTTCAAGCAACCAAGT 347
    |||
Db 181 TGAGCGTATGCTTATATTTGGCTAATGAGCTCAAAAAGATGAGACTTCAAGCAACCAAGT 240
QY 348 TCTTATGAAAGTTGCTTGAACAAGACATCAACGAGCTGCTTCAACGCTGCAAGCAAGCTTA 407
    |||
Db 241 TCTTATGAAAGTTGCTTGAACAAGACATCAACGAGCTGCTTCAACGCTGCAAGCAAGCTTA 300
QY 408 CCATGCTGCTACAGAAGTCTCTTGAAGAGAGAGCTGTGTACACACATACCGGTGAGCTT 467
    |||
Db 301 CCATGCTGCTACAGAAGTCTCTTGAAGAGAGAGCTGTGTACACACATACCGGTGAGCTT 360
```

Qy 468 CAGAAAGCTTTGGTTCTCTGTTACTCTATACAGATGAGAGATGAGTGAACAT 527  
 Db 361 CAGAAAGCTTTGGTTCTCTGTTACTCTATACAGATGAGAGATGAGTGAACAT 420  
 Qy 528 GACATTTGGCTTAAGCAAGAGCTTAAGCTGTCATGAGAAAATCAAGACAAGACTCA 587  
 Db 421 GACATTTGGCTTAAGCAAGAGCTTAAGCTGTCATGAGAAAATCAAGACAAGACTCA 480  
 Qy 588 TGAATGAGATGTTATTAAGATCTTGTCCAGAAAGCAAACTGATGATGATGATCTT 647  
 Db 481 TGAATGAGATGTTATTAAGATCTTGTCCAGAAAGCAAACTGATGATGATGATCTT 540  
 Qy 648 TAAACGTTACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 707  
 Db 541 TAAACGTTACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
 Qy 708 TGAATGAGATGTTCTGTCATTTGAGGTCAACATTCAGTCTTGAACAGACAGACT 767  
 Db 601 TGAATGAGATGTTCTGTCATTTGAGGTCAACATTCAGTCTTGAACAGACAGACT 660  
 Qy 768 TTAATTTGTCATGTTCTTCTGTCATGATCAACAAAATGAACTGATGAGAGACT 827  
 Db 661 TTAATTTGTCATGTTCTTCTGTCATGATCAACAAAATGAACTGATGAGAGACT 720  
 Qy 828 CACTAGAATTGTGACCAAGAGCTGAGATGATGATGATGATGATGATGATGATGAT 887  
 Db 721 CACTAGAATTGTGACCAAGAGCTGAGATGATGATGATGATGATGATGATGATGAT 780  
 Qy 888 GCGGAGAGAGAGATTCCTTTGAGAAAGCTATTACCAAGACAATGAGATGATGAT 947  
 Db 781 GCGGAGAGAGAGATTCCTTTGAGAAAGCTATTACCAAGACAATGAGATGATGAT 840  
 Qy 948 GAAATGCTGTCGTCATCTTCTGTCATGATGATGATGATGATGATGATGATGAT 1007  
 Db 841 GAAATGCTGTCGTCATCTTCTGTCATGATGATGATGATGATGATGATGATGAT 900  
 Qy 1008 AAACATAGAGCTGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1067  
 Db 901 AAACATAGAGCTGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 Qy 1068 GTTTCATATGTTGATTTGTTTCTACAAAACCTGTTGTTGTTGTTGTTGTTG 1127  
 Db 961 GTTTCATATGTTGATTTGTTTCTACAAAACCTGTTGTTGTTGTTGTTGTTG 1020  
 Qy 1128 AGTTCCTTAATTAAGCAAGAGAGA 1153  
 Db 1021 AGTTCCTTAATTAAGCAAGAGAGA 1046

RESULT 5  
 Locus CNSABV3 966 bp mRNA linear HTC 06-FEB-2004  
 DEFINITION Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 G5LTPGH18ZH06 of Hormone Treated Callus of strain col-0 of  
 Arabidopsis thaliana (thale cress).  
 Accession BX815939.1 GI:42472187  
 VERSION 1  
 SOURCE Arabidopsis thaliana (thale cress)  
 KEYWORDS HTC; G5LTP cDNA.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosoid II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 966)  
 Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,  
 Menard,M., Cruaud,C., Queirer,F., Scarpelli,C., Schachter,V.,  
 Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
 Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome  
 Annotation  
 Unpublished  
 JOURNAL 2 (bases 1 to 966)  
 REFERENCE  
 AUTHORS Genoscope.

TITLE Direct Submission  
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqret@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 COMMENT The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli  
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for Protein  
 Sequences) . 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
 length  
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
 FEATURES  
 Location/Qualifiers  
 1..966  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="G5LTPGH18ZH06"  
 /issue\_type="Hormone Treated Callus"  
 /plasmid="pCMVSPORT\_6"  
 1..966  
 /gene="At1g35720"  
 ORIGIN  
 Query Match 77.0%; Score 947.6; DB 3; Length 966;  
 Best Local Similarity 99.5%; Pred. No. 1,4e-236;  
 Matches 961; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
 Qy 195 CGAAACCTACGGGAGAGACCTTCTCAAGACTTTGACCAAGAGCTCTCAAGATTTGCA 254  
 Db 1 CGAGACTACGGGAGAGACCTTCTCAAGACTTTGACCAAGAGCTCTCAAGATTTGCA 60  
 Qy 255 GAGAGCTATCTTGTGTGACCTTTGAAACCGGTGAGCGTATGCTTATTTGGCTAATGA 314  
 Db 61 GAGAGCTATCTTGTGTGACCTTTGAAACCGGTGAGCGTATGCTTATTTGGCTAATGA 120  
 Qy 315 AGCTACAAAAGATGACCTTCAAGCAACAGTTTATGAGAGTCTTGGCAAGAGAC 374  
 Db 121 AGCTACAAAAGATGACCTTCAAGCAACAGTTTATGAGAGTCTTGGCAAGAGAC 180  
 Qy 375 ATCAAGCAGCTGCTTCAAGCTAGGAGCTTCAATGCTCGCTACCAAGAGTCTTGA 434  
 Db 181 ATCAAGCAGCTGCTTCAAGCTAGGAGCTTCAATGCTCGCTACCAAGAGTCTTGA 240  
 Qy 435 AGAGAGCTTCTACACACTACCGGTGACTTCAAGAACTTTGGTTCTCTGTATAC 494  
 Db 241 AGAGAGCTTCTACACACTACCGGTGACTTCAAGAACTTTGGTTCTCTGTGTATC 300  
 Qy 495 CTGATACAGGTACGAAGAGATGAATGAACATGATGATGATGATGATGATGATGAT 554  
 Db 301 CTGATACAGGTACGAAGAGATGAATGAACATGATGATGATGATGATGATGATGAT 360  
 Qy 555 GGTCCATGAGAAATCAAGGACCAAGCACTCAATGATGATGATGATGATGATGATGAT 614  
 Db 361 GGTCCATGAGAAATCAAGGACCAAGCACTCAATGATGATGATGATGATGATGATGAT 420  
 Qy 615 CACAAGAGCAAGCTCAATCAATCTTAAACCGTTACCAAGATGATGATGATGATGAT 674  
 Db 421 CACAAGAGCAAGCTCAATCAATCTTAAACCGTTACCAAGATGATGATGATGATGAT 480  
 Qy 675 GGAATTTCCAAAGCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGAT 734  
 Db 481 GGAATTTCCAAAGCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGAT 540  
 Qy 735 GTCAACATTCAGTGTGATCAAGACCAAGACTTTTCTGTCATGTTCTTCTGTTCACT 794  
 Db 541 GTCAACATTCAGTGTGATCAAGACCAAGACTTTTCTGTCATGTTCTTCTGTTCACT 600

QY	795	AATCAAAACAACTGGAACGTAN	GAAGAGC	CTCA	CTGAA	TTTGAC	CAACAAG	CTGA	854
Db	601	AATCAACAAACTGGAACTGAT	GAAGAGCA <th>CTCA</th> <td>CTGAA</td> <td>TTTGAC <td>CAACAAG <td>CTGA</td> <td>660</td> </td></td>	CTCA	CTGAA	TTTGAC <td>CAACAAG <td>CTGA</td> <td>660</td> </td>	CAACAAG <td>CTGA</td> <td>660</td>	CTGA	660
QY	855	GATTGACTTGAAGGCTCAT	TGAGAGAG <td>AGTAC</td> <td>CGCAG</td> <td>GAACAG</td> <td>CAGAT</td> <td>TCCTTTGGANA</td> <td>914</td>	AGTAC	CGCAG	GAACAG	CAGAT	TCCTTTGGANA	914
Db	661	GATTGACTTGAAGGCTCAT	TGAGAGAG <td>AGTAC</td> <td>CGCAG</td> <td>GAACAG</td> <td>CAGAT</td> <td>TCCTTTGGANA</td> <td>720</td>	AGTAC	CGCAG	GAACAG	CAGAT	TCCTTTGGANA	720
QY	915	AGCTATTACCAAAAGCACT	CGTGGAG <td>ATTAC</td> <td>GAAAG <td>TGCTCG <td>TCGCA <td>TTTCGGTGA</td> <td>974</td> </td></td></td>	ATTAC	GAAAG <td>TGCTCG <td>TCGCA <td>TTTCGGTGA</td> <td>974</td> </td></td>	TGCTCG <td>TCGCA <td>TTTCGGTGA</td> <td>974</td> </td>	TCGCA <td>TTTCGGTGA</td> <td>974</td>	TTTCGGTGA	974
Db	721	AGCTATTACCAAAAGCACT	CGTGGAG <td>ATTAC</td> <td>GAAAG <td>TGCTCG <td>TCGCA <td>TTTCGGTGA</td> <td>780</td> </td></td></td>	ATTAC	GAAAG <td>TGCTCG <td>TCGCA <td>TTTCGGTGA</td> <td>780</td> </td></td>	TGCTCG <td>TCGCA <td>TTTCGGTGA</td> <td>780</td> </td>	TCGCA <td>TTTCGGTGA</td> <td>780</td>	TTTCGGTGA	780
QY	975	AGATGATGCTTAATCAAT	CAATCC <td>CCACAG</td> <td>AGAAAC</td> <td>TTAAG</td> <td>CTCT</td> <td>CTACAGCTTCTGT</td> <td>1034</td>	CCACAG	AGAAAC	TTAAG	CTCT	CTACAGCTTCTGT	1034
Db	781	AGATGATGCTTAATCAAT	CAATCC <td>CCACAG</td> <td>AGAAAC</td> <td>TTAAG</td> <td>CTCT</td> <td>CTACAGCTTCTGT</td> <td>840</td>	CCACAG	AGAAAC	TTAAG	CTCT	CTACAGCTTCTGT	840
QY	1035	TATCTCTATCTCCCTCT	CTCTCT	CTTTG	ATGATG	ATTTCAAA	TGCTTTG	ATTTGTTTCTA	1094
Db	841	TATCTCTATCTCCCTCT	CTCTCT	CTTTG	ATGATG	ATTTCAAA	TGCTTTG	ATTTGTTTCTA	900
QY	1055	CAAAA-ACCTTGTTGTTT	CTGTGTG <td>TTTGAG</td> <td>TTTGAG</td> <td>TTTGAG</td> <td>TTTGAG</td> <td>TTTGAG</td> <td>1155</td>	TTTGAG	TTTGAG	TTTGAG	TTTGAG	TTTGAG	1155
Db	901	CAAAAACACCTTGTTGTT	CTGTGTG <td>TTTGAG</td> <td>TTTGAG</td> <td>TTTGAG</td> <td>TTTGAG</td> <td>TTTGAG</td> <td>960</td>	TTTGAG	TTTGAG	TTTGAG	TTTGAG	TTTGAG	960
QY	1154	GACAGA	1159						
Db	961	GAAAGA	966						
RESULT 6	CNS09Y3M	1034 bp	mRNA	linear	HTC 06-FEB-2004				
LOCUS	CNS09Y3M	1034 bp	mRNA	linear	HTC 06-FEB-2004				
DEFINITION	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSU783JZ801 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).								
ACCESSION	CNS09Y3M	1034 bp	mRNA	linear	HTC 06-FEB-2004				
VERSION	1								
KEYWORDS	Arabidopsis thaliana (thale cress)								
SOURCE	Arabidopsis thaliana (thale cress)								
ORGANISM	Arabidopsis thaliana (thale cress)								
REFERENCE	Castelli,V., Aury,J.M., Jallion,O., Winkler,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat M. Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation								
TITLE	Unpublished								
JOURNAL	Genoscope.								
REFERENCE	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr								
AUTHORS	Web : www.genoscope.cns.fr								
JOURNAL	The sequences are based on single pass reads.								
COMMENT	Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Winkler P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Ptipap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Fulllength								
FEATURES	http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis								
SOURCE	Location/Qualifiers								
	1..1034								

		/organism="Arabidopsis thaliana"
		/mol_type="mRNA"
		/strain="Col-0"
		/db_xref="taxon:3702"
		/clone="GSLTIS23ZB01"
		/tissue_type="Adult vegetative tissue"
		/plasmid="PCWMSport_6"
		1..1034
		/gene="Atlg35720"
ORIGIN		
Query Match	75.1%; Score 924.2; DB 3;	Length 1034;
Best Local Similarity	98.0%; Pred. No. 1,9e-230;	
Matches 946;	Conservative 0;	Mismatches 18; Indels 1; Gaps 1
OY	12 AAACCTAAAAGTAGAAGAAAAATGGCACTCTTAAGCTTTTCGTATTCGTTCCTGCCTCC	71
DB	1 AAAACCTAAAAGTAGAAGAAATGGCACTCTTAAGCTTTTCGTATTCGTTCCTGCCTCC	60
OY	72 TTCGTATGATGCTGAACAATTGAGAACCGCTTTTGAGAATGGGGTACCAAGAGACTT	131
DB	61 TTCGTATGATGCTGAACAATTGAGAACCGCTTTTGAGAATGGGGTACCAAGAGACTT	120
OY	132 GATCATATCATCTTGGCTCACAGAAAGTCTGAAACAGAGAAAGTCATCAGGCAAGCAT	191
DB	121 GATCATATCATCTTGGCTCACAGAAAGTCTGAAACAGAGAAAGTCATCAGGCAAGCAT	180
OY	192 CCACGAAACCTACCGCGCAAGACCTTCTCAAGACTCTTGACAGAGAGCTCTCAAGATTT	251
DB	181 CCAAGAGACTACCGCGCAAGACCTTCTCAAGACTCTTGACAGAGAGCTCTCAAGATTT	240
OY	252 CGAGAGAGCTATCTTGTGTGGAGCTCTTGAACCCGGTGAAGCTGATGCTTTATTTGGCTAA	311
DB	241 CGAGAGAGCTATCTTGTGTGGAGCTCTTGAACCCGGTGAAGCTGATGCTTTATTTGGCTAA	300
OY	312 TGAAGCTACAAAAAGATGSACTTCAAGCAACCAAGTTCCTATGSAAGTTCCTTGACAAG	371
DB	301 TGAAGCTACAAAAAGATGSACTTCAAGCAACCAAGTTCCTATGSAAGTTCCTTGACAAG	360
OY	372 GACATCAACGAGAGCTGCTTCAACGCTAGGCAAGCTTACATGCTGCTCAACAAGAGCTCT	431
DB	361 GACATCAACGAGAGCTGCTTCAACGCTAGGCAAGCTTACATGCTGCTCAACAAGAGCTCT	420
OY	432 TGAAGAGGAGCTTGCTCAACAACCTACCGGTGACTTCAAGAAAGCTTTTGCTTCTGCT	491
DB	421 TGAAGAGGAGCTTGCTTGCTCAACAACCTACCGGTGACTTCAAGAAAGCTTTTGCTTCTG	480
OY	492 TAACCTCAACAGGTACGAAAGAGATGAAGTGAACATGCAATGGCTAAGCAAGAGCTAA	551
DB	481 TAACCTCAACAGGTACGAAAGAGATGAAGTGAACATGCAATGGCTAAGCAAGAGCTAA	540
OY	552 GCTGGTCATGAGAAATCAAGSACAAACACTACATCAATGATGAGATGTTATTAAGATCTT	611
DB	541 GCTGGTCATGAGAAATCAAGSACAAACACTACATCAATGATGAGATGTTATTAAGATCTT	600
OY	612 GTCCAACAAGAACCAAGCTCAGATCATGTCTTCTTAAACGCTTACCAAAGATGATCATGG	671
DB	601 GTCCAACAAGAACCAAGCTCAGATCATGTCTTCTTAAACGCTTACCAAAGATGATCATGG	660
OY	672 CGAGGAATTCCTCAAGAGCTTTGAGSAGAGGATGATGATGATGATGATGATGATGATGATG	731
DB	661 CGAGGAATTCCTCAAGAGCTTTGAGSAGAGGATGATGATGATGATGATGATGATGATGATG	720
OY	732 GAGGTCAACATTCAGTCTTGAACAAGACAGAGCTTATCTTGTGATGATGATGATGATGATG	791
DB	721 GAGGTCAACATTCAGTCTTGAACAAGACAGAGCTTATCTTGTGATGATGATGATGATGATG	780
OY	792 AGCAATCAACAAAACCTGGAATGATGAAGAGACTCACTGAATTTGTGACCAACAAGAC	851
DB	781 AGCAATCAACAAAACCTGGAATGATGAAGAGACTCACTGAATTTGTGACCAACAAGAC	840
OY	852 TGAAGTTGACTTGAAGTTCATTTGAGAGAGAGTACAGGCGACAGAAACAGATTCCTTTGGA	911



Db 841 TGAGATTGACTTGAAAGTTCATTGGAGAGATACAGGACGAGAAACGACTTCCTTTGA 900

QY 912 GAAAGCTATTACCAAGACACTCGTGGAGATTAGAGAAAGTCTCGTGGACTTCGG 971

Db 901 GAGAGCTATTACCAAGACACTCGTGGAGATTAGAGAGGA-GCTCGTGGACTTCGGA 959

QY 972 TGAAG 976

Db 960 TGAGG 964

RESULT 7  
CF652924 910 bp mRNA linear EST 06-NOV-2003  
LOCUS 85-L020135w-066-001-121-SP6P MP1Z-ADIS-066 Arabidopsis thaliana  
DEFINITION cDNA clone MP1Z20011211Q 5-PRIME, mRNA sequence.  
ACCESSION CF652924  
VERSION CF652924.1 GI:37429878  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Bukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 910)  
Schmid,K.U., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,  
Mitchell-Olds,T. and Weisshaar,B.  
Large-scale identification and analysis of genome-wide  
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana  
Genome Res. 13 (6), 1250-1257 (2003)  
22683290  
12799357  
Contact: Weisshaar B  
ADIS DNA core facility at MP1Z  
Max-Planck-Institute for Plant Breeding Research  
Carl-von-Linne Weg 10, 50829 Koeln, Germany  
Fax: 00492215062851  
Email: weisshaar@mpiz-koeln.mpg.de  
Insert Length: 910 Std Error: 0.00  
Plate: 1 row: 1 column: 21  
Seq Primer: SP6P.  
Location/Qualifiers  
1. 910  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Ws-0"  
/db\_xref="GABI:939529"  
/db\_xref="taxon:3702"  
/clone="MP1Z20011211Q"  
/issue\_type="root"  
/lab\_host="E. coli TOP10"  
/clone\_idb="MP1Z-ADIS-066"  
/note="Vector: PCMVSPORT6; Site 1: SalI; Site 2: NotI;  
cDNA library from Arabidopsis thaliana, accession  
Masilewskija-0; roots from three weeks old plants grown  
on MS-plates at 26M-0C with 16 hours light/day; library  
was made at the Max-Planck-Institute for Plant Breeding  
Research, Cologne, Germany; cloning sites SalI-NotI,  
primer sites and orientation:  
SP6-Sali-CCACGCGCG-5prime-cDNA-polyA-CC-NotI-T7; GATEWAY  
compatible; Note: Sequencing granted in the context of the  
GABI Arabidopsis Verbund I; Genetic Diversity,  
'establishment of high-efficiency SNP-based mapping tools  
and development of methods for genome-wide mutation  
detection'; PI: Bernd Weisshaar Sequence submission managed  
by RZPD/GABI-Primary database: <http://gabi.rzpd.de> This  
clone is available from RZPD; contact RZPD ([clone@rzpd.de](mailto:clone@rzpd.de))  
for further information."

ORIGIN  
Query Match 72.5%; Score 892.2; DB 7; Length 910;  
Best Local Similarity 99.0%; Pred. No. 4,3e-222;  
Matches 897; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 21 AAGTAGAAGAAAATGGCGACTCTTAAGGTTTCTGATTCTTCTGCTCTTCTGATGA 80

Db 5 AAGTAGAAGAAAATGGCGACTCTTAAGGTTTCTGATTCTTCTGCTCTTCTGATGA 64

QY 81 TGTGAGCAATTGAGAACCGCTTTTGAAGGATGGGGTACGAAAGAGGACTTGATCATATC 140

Db 65 TGTGAGCAATTGAGAACCGCTTTTGAAGGATGGGGTACGAAAGAGGACTTGATCATATC 124

QY 141 AATCTTGCTCAGAGAGTCTGAGCAAGAGAAAGTCATAGGCAAGCATACGAAAC 200

Db 125 AATCTTGCTCAGAGAGTCTGAGCAAGAGAAAGTCATAGGCAAGCATACGAAAC 184

QY 201 CTACGCGAAGACCTTCTCAAGACTCTTGACAAAGAGCTCTTAAGATTTCGAGAGC 260

Db 185 CTACGCGAAGACCTTCTCAAGACTCTTGACAAAGAGCTCTTAAGATTTCGAGAGC 244

QY 261 TATCTTGTTGTGACTCTTGAACCGGATGAGGCTTATGAGCTTAATGAGCTAC 320

Db 245 TATCTTGTTGTGACTCTTGAACCGGATGAGGCTTATGAGCTTAATGAGCTAC 304

QY 321 AAAAAGATGACTTCAAGCAACCAAGTTCTATGAAAGTGTGCAAGAGCATCAAC 380

Db 305 AAAAAGATGACTTCAAGCAACCAAGTTCTATGAAAGTGTGCAAGAGCATCAAC 364

QY 381 GCACTGCTTCAAGCTAGGCAAGCTTACCATGCTCCCTACAGAAAGTCTTGAAGAG 440

Db 365 GCACTGCTTCAAGCTAGGCAAGCTTACCATGCTCCCTACAGAAAGTCTTGAAGAG 424

QY 441 CGTTGCTACCACTACCGGTGACTTCAAGAAAGCTTTTGTTCTTGTATTCCTAT 500

Db 425 CGTTGCTACCACTACCGGTGACTTCAAGAAAGCTTTTGTTCTTGTATTCCTAT 484

QY 501 CAGGTACGAAGAGATGAAGTGAATGACATTTGGCTTAAGCAAGAGCTAAGCTGTCA 560

Db 485 CAGGTACGAAGAGATGAAGTGAATGACATTTGGCTTAAGCAAGAGCTAAGCTGTCA 544

QY 561 TGAAGAAATCAGAGCAAGACATCAATGATGAGATGTTATTGAATCTTGTCCACAG 620

Db 545 TGAAGAAATCAGAGCAAGACATCAATGATGAGATGTTATTGAATCTTGTCCACAG 604

QY 621 AACCAAGCTCAGATCAATGCTATCTTTAACGTTTCAAGATGATCAAGGAGAAAT 680

Db 605 AACCAAGCTCAGATCAATGCTATCTTTAACGTTTCAAGATGATCAAGGAGAAAT 664

QY 681 TCTCAAGAGCTTGAAGAGAGATGATGATGACAGTTCTTGCATTTTGAAGTCAAC 740

Db 665 TCTCAAGAGCTTGAAGAGAGATGATGATGACAGTTCTTGTGTGAGTCAAC 724

QY 741 CATTCAAGTCTTGAAGAGACCAAGAGCTTTTGTGATGTTCTTGTCAAGCAATCA 800

Db 725 CATTCAAGTCTTGAAGAGACCAAGAGCTTTTGTGATGTTCTTGTCAAGCAATCA 784

QY 801 CAAACATGAATCAGAGAGAGCACTCACTGAATTGACCAAGAGCTGAGTTGA 860

Db 785 CAAACATGAATCAGAGAGAGCACTCACTGAATTGACCAAGAGCTGAGTTGA 844

QY 861 CTTGAAGCTATTGAGAGAGATGATCAGGCGAGAAACAGATCTCTTGAAGAACTAT 920

Db 845 CTTGAAGCTATTGAGAGAGATGATCAGGCGAGAAACAGATCTCTTGAAGAACTAT 904

QY 921 TACCAA 926

Db 905 TACCAA 910

RESULT 8  
LOCUS CDSOAC11  
DEFINITION CDSOAC11 865 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLF692F06 of Flowers and buds of strain col-0 of Arabidopsis  
thaliana (thale cress).  
ACCESSION EX814350

```

VERSION      BX814350.1  GI:42472041
KEYWORDS     HTC; GSUT cDNA.
SOURCE       Arabidopsis thaliana (thale cress)
ORGANISM     Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE    1 (bases 1 to 865)
AUTHORS      Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
              Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
              Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE        Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
              A Combined Approach to Evaluate and Improve Arabidopsis Genome
              Annotation
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 865)
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT      The sequences are based on single pass reads.
              Life Technologies (a division of Invitrogen) members carried out
              full-length libraries construction : Temple G.
              Genoscope members carried out sequencing and annotation : Castelli
              V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
              Schachter V., Weissenbach J., Salanoubat M.
              URGV INRA : Clepet C., Caboche M.
              Annotation is based on the June 2003 version of the Arabidopsis
              genome released by MIPS (Munich Information center for Protein
              Sequences) . 5 prime and 3 prime are assembled with Phrap.
              http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
              length
              http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
              location/Qualifiers
FEATURES
  source
    1..865
    /organism="Arabidopsis thaliana"
    /mol_type="mRNA"
    /strain="Col-0"
    /db_xref="taxon:3702"
    /clone="GSUTRB69ZP06"
    /tissue_type="Flowers and buds"
    /plasmid="PCMVSPORT_6"
    1..865
    /gene="At1g35720"
ORIGIN
Query Match      68.0%; Score 836.2; DB 3; Length 865;
Best Local Similarity 97.9%; Pred. No. 1.9e-207;
Matches 847; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 349 CTTATGGAAGTTCCTGCAACAAGACATCAACGACGCTGTTACGCTAGGCAAGCTTAC 408
DB 1 CTTATGGAAGTTCCTGCAACAAGACATCAACGACGCTGTTACGCTAGGCAAGCTTAC 60
QY 409 CATGCTCGCTACAAGAAGTCTCTTGAAGAAGAGCTTGCTACACACACTACCGTGACTTC 468
DB 61 CATGCTCGCTACAAGAAGTCTCTTGAAGAAGAGCTTGCTACACACACTACCGTGACTTC 120
QY 469 AGAAAGCTTTGGTTCTCTTGTTCCTCATACAGGTACGAGAGAGATGAAGTGAACATG 528
DB 121 AGAAAGCTTTGGTTCTCTTGTTCCTCATACAGGTACGAGAGATGAAGTGAACATG 180
QY 529 ACATTTGGCTTAAGAAGAAGTACGTGCTGCATGAGAAAATCAAGAACAAGCACTACAT 588
DB 181 ACATTTGGCTTAAGAAGAAGTACGTGCTGCATGAGAAAATCAAGAACAAGCACTACAT 240
QY 589 GATGAGAGATGTTATTAAGATCTTGCTCAACAAGAAGCAAGCTCAGATCATGCTACTTTT 648
DB 241 GATGAGAGATGTTATTAAGATCTTGCTCAACAAGAAGCAAGCTCAGATCATGCTACTTTT 300
QY 649 AACCGTTTACCAAGATGATCATGCGCAGGAATTTCTCAAGAGTCTTAGAGAGAGATGAT 708

```

```

DB 301 AACCGTTTACCAAGATGATCATGCGCAGGAATTTCTCAAGAGTCTTAGAGAGAGATGAT 360
QY 709 GATGACAAAGTTCCTTGCACTTTTGAAGGTCAACATTCAGTGTGTAACAAGACAGCTT 768
DB 361 GATGACAAAGTTCCTTGCACTTTTGAAGGTCAACATTCAGTGTGTAACAAGACAGCTT 420
QY 769 TACTTTGTGATGTTCTTGTGTTTGTTCAGCAATCAACAAATGCAAGTGAAGAGACACTC 828
DB 421 TACTTTGTGATGTTCTTGTGTTTGTTCAGCAATCAACAAATGCAAGTGAAGAGACACTC 480
QY 829 ACTAGAAATGTGACCAACAAGACCTGAGATTTGACTTGAAGGTATTTGAGAGAGATTCAG 888
DB 481 ACTAGAAATGTGACCAACAAGACCTGAGATTTGACTTGAAGGTATTTGAGAGAGATTCAG 540
QY 889 CCGAGGAACAGCAATTCCTTTGAGAAAGCTATTAACAAGACACTGTGTGAGATTACGAG 948
DB 541 TCGAGGAACAGCAATTCCTTTGAGAAAGCTATTAACAAGACACTGTGTGAGATTACGAG 600
QY 949 AAGATGCTGTGCGCACTTCTCGGTGAAGATGATGCTTAATCAATCAATCTCCACAGAGA 1008
DB 601 AAGATGCTGTGCGCACTTCTCGGTGAAGATGATGCTTAATCAATCAATCTCCACAGAGA 660
QY 1009 AACATTAAGTGTCTTACACAGCTTCTGTATCTTATCTCCCTCTCTCTCTTTGATGAG 1068
DB 661 AACATTAAGTGTCTTACACAGCTTCTGTATCTTATCTCCCTCTCTCTCTTTGATGAG 720
QY 1069 TTTCAATGCTTGAATTTGTTTCTTACAAAACCTGTTGTTGTTGTTGTTGTTTGA 1128
DB 721 TTTCAATGCTTGAATTTGTTTCTTACAAAACCTGTTGTTGTTGTTGTTGTTTGA 780
QY 1129 GTTCTTAATTAATGCAAAAGAGAGACAGAGACACAGTGTGCTCTTAAGTTAAT 1188
DB 781 GTTCTTAATTAATGCAAAAGAGAGACAGAGACACAGTGTGCTCTTAAGTTAAT 840
QY 1189 ATATATGAAGACATTTGGCTTAAA 1213
DB 841 ATATATGAAGACATTTGGCTTAAA 865

RESULT 9
LOCUS       CF652048                      843 bp      mRNA      linear      EST 06-NOV-2003
DEFINITION  30-1020580-066-004-L08-SP6P MP1Z-ADIS-066 Arabidopsis thaliana cDNA
            Clone MP1Zp2001L084Q 5-PRIME, mRNA sequence.
ACCESSION   CF652048
VERSION     CF652048.1  GI:37428171
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE    1 (bases 1 to 843)
AUTHORS      Schmidt,K.U., Soeensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
              Mitchell-Olds,T. and Weisshaar,B.
TITLE        Large-scale identification and analysis of genome-wide
              single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
JOURNAL      Genome Res. 13 (6), 1250-1257 (2003)
MEDLINE     22683290
PUBMED      12799357
COMMENT      Contact: Weisshaar B
              ADIS DNA core facility at MP1Z
              Max-Planck-Institute for Plant Breeding Research
              Carl-von-Linne Weg 10, 50829 Koeln, Germany
              Fax: 00492215062851
              Email: weisshaar@mpiz-koeln.mpg.de
              Insert length: 843 Std Error: 0.00
              Place: 4 row: 1 column: 08
              Seq primer: SP6P.
FEATURES
  source
    1..843
    /organism="Arabidopsis thaliana"
    /mol_type="mRNA"

```

/ecotype="Ws-0"  
/db\_xref="GABI:938637"  
/db\_xref="taxon:3702"  
/clone="MP12001L084Q"  
/issue\_type="root"  
/lab\_host="E. coli TOP10"  
/clone\_11b="MP12-ADIS-066"  
/note="Vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI;  
cDNA library from Arabidopsis thaliana, accession  
Wassilewskija-0; roots from three weeks old plants grown  
on MS-plates at 26M-0C with 16 hours light/day; library  
was made at the Max-Planck-Institute for Plant Breeding  
Research, Cologne, Germany; cloning sites SalI-NotI,  
primer sites and orientation:  
SP6-Sali-CCACCGCTCCG-5prime-cDNA-polyA-CC-NotI-T7, GATEWAY  
compatible! Note: Sequencing granted in the context of the  
GABI Arabidopsis Verbund I: Genetic Diversity,  
'Establishment of high-efficiency SNP-based mapping tools  
and development of methods for genome-wide mutation  
detection', PI: Bernd Weisshaar Sequence submission managed  
by RZPD/GABI-Primary database: <http://gabi.rzpd.de> This  
clone is available from RZPD; contact RZPD (clone@rzpd.de)  
for further information."

## ORIGIN

Query Match 67.6%; Score 831.8; DB 7; Length 843;  
Best Local Similarity 99.2%; Pred. No. 2.7e-206;  
Matches 836; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 21 AAGTAGAAGAAAATGCGACTCTTAAGGTTCTGATTCCTGCTCTTCTGATGA 80  
DB 1 AAGTAGAAGAAAATGCGACTCTTAAGGTTCTGATTCCTGCTCTTCTGATGA 60

QY 81 TCGTAGCAATTGAAACCGCTTTTGAAGATGCGGTACGAAAGAGAGAGAGAGAG 140  
DB 61 TGCTGACCAATTGAAACCGCTTTTGAAGATGCGGTACGAAAGAGAGAGAGAG 120

QY 141 AATCTTGAGCTCAGAGAGTGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 200  
DB 121 AATCTTGAGCTCAGAGAGTGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180

QY 201 CTACGCGAG 260  
DB 181 CTACGCGAG 240

QY 261 TATCTTGTGTGACTCTTGAACCGGTGAGCGGTGATGCTTTATTTGCTAATGAAGCTAC 320  
DB 241 TATCTTGTGTGACTCTTGAACCGGTGAGCGGTGATGCTTTATTTGCTAATGAAGCTAC 300

QY 321 AAAAAGATGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 380  
DB 301 AAAAAGATGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 381 GCAAGCTCTTCAAG 440  
DB 361 GCAAGCTCTTCAAG 420

QY 441 CGTTGCTACCACTACCGGTGACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 500  
DB 421 CGTTGCTACCACTACCGGTGACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480

QY 501 CAGGTAG 560  
DB 481 CAGGTAG 540

QY 561 TGAAGAAATCAAG 620  
DB 541 TGAAGAAATCAAG 600

QY 621 AAGCAAGAGCTAGATGATCTTTTAAACCGTTACCAAGAGAGAGAGAGAGAGAGAG 680  
DB 601 AAGCAAGAGCTAGATGATCTTTTAAACCGTTACCAAGAGAGAGAGAGAGAGAGAG 660

QY 681 TCTCAAGAGCTTGAAG 740  
DB 661 TCTCAAGAGCTTGAAG 720

QY 741 CATTCAAGCTTGAAG 800  
DB 721 CATTCAAGCTTGAAG 780

QY 801 CAAAACCTGAAG 860  
DB 781 CAAAACCTGAAG 840

QY 861 CTT 863  
DB 841 CTT 843

RESULT 10  
LOCUS CA782049 814 bp mRNA linear EST 03-DEC-2002  
DEFINITION O1E11AF Infected Arabidopsis leaf Arabidopsis thaliana cDNA, mRNA  
sequence.  
ACCESSION CA782049  
VERSION CA782049.1 GI:26020095  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 814)  
Lundgaard,M., Emeersen,J., Nielsen,K.L., Wilson,I., Somerville,S.  
and Wellinder,K.G.  
EST sequencing of Erysiphe cichoracearum infected Arabidopsis  
plants  
TITLE Unpublished (2002)  
JOURNAL Contact: Karen G. Wellinder  
COMMENT Institut for bioteknologi  
Aalborg Universitet  
Sohngaardsholmvej 49, 9000 Aalborg, Denmark  
Tel: +45 96358467  
Fax: +45 98141808  
Email: kgeb@bio.auc.dk.  
FEATURES  
Location/Qualifiers  
source  
1..814  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Colombia"  
/db\_xref="taxon:3702"  
/dev\_stage="Plant 3 weeks old, three days post infection"  
/clone\_11b="Infected Arabidopsis leaf"  
/note="Organ: Leaf; Vector: pBluescript; Mixed cDNA  
library of Arabidopsis and E. cichoracearum infected leaf  
from three weeks old Arabidopsis plants were  
harvested 3 days after infection and mRNA oligo dT  
selected."

ORIGIN

Query Match 63.9%; Score 786.2; DB 6; Length 814;  
Best Local Similarity 99.4%; Pred. No. 2.3e-194;  
Matches 810; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 28 AGAAAATGCGAGCTTGAAGTTCTGATTCCTGCTCTTCTGATGATGCTGAG 87  
DB 1 AGAAAATGCGAGCTT-AGGTTCTGATTCGTTCTGCTCTTCTGATGATGCTGAG 59

QY 88 CAATTGGAACCGCTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 147  
DB 60 CAATTGGAACCGCTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 119

QY 148 GCTCAGAGAGTGTGAG 207  
DB 120 GCTCAGAGAGTGTGAG 179

```
QY 208 GAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTAACGATTTGAGAGAGCTATCTG 267
    |||
Db 180 GAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTAACGATTTGAGAGAGCTATCTG 239
QY 268 TTGTGACTCTTGAACCCGGTGAAGCTGATGCTTTATTTGGCTAATGAAGCTACAAAAGA 327
    |||
Db 240 TTGTGACTCTTGAACCCGGTGAAGCTGATGCTTTATTTGGCTAATGAAGCTACAAAAGA 299
QY 328 TGGACTTCAAGCAACCAAGTTCTTATGGAAGTTGCTTGCAAGAGACATCAAGCAGCTG 387
    |||
Db 300 TGGACTTCAAGCAACCAAGTTCTTATGGAAGTTGCTTGCAAGAGACATCAAGCAGCTG 359
QY 388 CTTCACGCTAAGGAGGATTAACATGCTCGCTCAAGAAGCTCTTAAGAAGAGAGCTTGC 447
    |||
Db 360 CTTCACGCTAAGGAGGATTAACATGCTCGCTCAAGAAGCTCTTAAGAAGAGAGCTTGC 419
QY 448 CACCACTAACCCTGATCTTCAAGAAAGCTTTGGTTCTCTTGTACCTCATACAGTAC 507
    |||
Db 420 CACCACTAACCCTGATCTTCAAGAAAGCTTTGGTTCTCTTGTACCTCATACAGTAC 479
QY 508 GAAGGAGATGAATGAACATGATGCTTGAAGCAAGAGCTAAGCTGCTCATGAGAA 567
    |||
Db 480 GAAGGAGATGAATGAACATGATGCTTGAAGCAAGAGCTAAGCTGCTCATGAGAA 539
QY 568 ATCAAGGACAAAGACTAATGATGAGATGATTATGATCTTGTCAAGAAGCAAA 627
    |||
Db 540 ATCAAGGACAAAGACTAATGATGAGATGATTATGATCTTGTGTCCAGAGACAAA 599
QY 628 GCTCAGATCAATGCTATCTTTAACCGTTACCAAGATGATCATGCGAGAAATTTCAAG 687
    |||
Db 600 GCTCAGATCAATGCTATCTTTAACCGTTACCAAGATGATCATGCGAGAAATTTCAAG 659
QY 688 AGTCTTGAAGAGAGATGATGATGACAAGTTCTTGCATTTTGAAGTCAACCATTCAG 747
    |||
Db 660 AGTCTTGAAGAGAGATGATGATGACAAGTTCTTGCATTTTGAAGTCAACCATTCAG 719
QY 748 TGCCTTGAACAAGACAGAGCTTTACTTTGATGATGTTCTTGCATCAAGATCAACAAACT 807
    |||
Db 720 TGCCTTGAACAAGACAGAGCTTTACTTTGATGATGTTCTTGCATCAAGATCAACAAACT 779
QY 808 GGAAGTG-ATGAAGAGACACTACTAGAAATTGTA 841
    |||
Db 780 GGAAGTGCTGAAGAGACACTACTAGAAATTGTA 814

RESULT 11
BU635010 782 bp mRNA linear EST 23-SEP-2002
LOCUS 003A04 Infected Arabidopsis Leaf Arabidopsis thaliana cdna, mRNA
DEFINITION
sequence.
ACCESSION BU635010
VERSION BU635010.1 GI:23302265
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 782)
AUTHORS Lundgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S.
and Welinder,K.G.
EST sequencing of Erysiphe cichoracearum infected Arabidopsis
plants
JOURNAL Unpublished (2002)
COMMENT Contact: Karen G. Welinder
Institute for biotechnology
Aalborg Universitet
Sohngaardsholmevej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.auc.dk.
FEATURES
Location/Qualifiers
```

```
source
1..782
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/dev_stage="Plant 3 weeks old, three days post infection"
/clone_lib="Infected Arabidopsis Leaf"
/notes="Organ: Leaf; Vector: pBluescript; Mixed cDNA
library of Arabidopsis and E. cichoracearum infected leaf
from three weeks old Arabidopsis plants. Plants were
harvested 3 days after infection and mRNA oligo dt
selected."

ORIGIN
Query Match 60.5%; Score 743.6; DB 5; Length 782;
Best Local Similarity 99.2%; Pred. No. 3,1e-183;
Matches 768; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY 26 GAAGAAAATGCGACTCTTAAAGTTCGATTCCTGCTCCTTCTGATGATGCTG 85
    |||
Db 11 GAGAGAAAATGCGACTCTTAAAG-TTCTGATTCG-TCTGCTCCTTCTGATGATGCTG 68
QY 86 AGCAATTGAAACCGCTTTTGAAGATGGGTTACGAAGAGACTTGAATCATATCAATCT 145
    |||
Db 69 AGCAATTGAAACCGCTTTTGAAGATGGGTTACGAAGAGACTTGAATCATATCAATCT 128
QY 146 TGGCTCAGAGAGTCTGAACAGAGAAAGTATCAGGCAAGCATACGAAACCTACG 205
    |||
Db 129 TGGCTCAGAGAGTCTGAACAGAGAAAGTATCAGGCAAGCATACGAAACCTACG 188
QY 206 GCGAAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTAACGATTTGAGAGAGCTATCT 265
    |||
Db 189 GCGAAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTAACGATTTGAGAGAGCTATCT 248
QY 266 TGTGTGAGCTTTGAACCCGGTGAAGCTGATGCTTTATTTGGCTAATGAAGTCAAAA 325
    |||
Db 249 TGTGTGAGCTTTGAACCCGGTGAAGCTGATGCTTTATTTGGCTAATGAAGTCAAAA 308
QY 326 GATGACTTCAAGCAACCAAGTCTTATGGAAGTGTGCTGCAACAAGACATCAAGCAGC 385
    |||
Db 309 GATGACTTCAAGCAACCAAGTCTTATGGAAGTGTGCTGCAACAAGACATCAAGCAGC 368
QY 386 TGCCTCAGCGTGAAGAGCTTACATGCTCGCTACAGAAGTCTCTTGAAGAGAGCTTG 445
    |||
Db 369 TGCCTCAGCGTGAAGAGCTTACATGCTCGCTACAGAAGTCTCTTGAAGAGAGCTTG 428
QY 446 CTCACCACTACCGGTGACTTCAGAAAGCTTTTGCTTCTTGTATCTATACAGGT 505
    |||
Db 429 CTCACCACTACCGGTGACTTCAGAAAGCTTTTGCTTCTTGTATCTATACAGGT 488
QY 506 AGGAGGAGATGAAGATGAACATGATGGCTTAAGCAAGAGCTAAGCTGCTCATGAGA 565
    |||
Db 489 AGGAGGAGATGAAGATGAACATGATGGCTTAAGCAAGAGCTAAGCTGCTCATGAGA 548
QY 566 AATCAAGGACAAAGACTAATGATGAGATGTTATGAAATCTTGTCCACAAGAGCA 625
    |||
Db 549 AATCAAGGACAAAGACTAATGATGAGATGTTATGAAATCTTGTCCACAAGAGCA 608
QY 626 AAGCTCAGATCATGCTTACTTTTAAACCGTTACCAAGATGATATGCGAGAGAAATTTCTCA 685
    |||
Db 609 AAGCTCAGATCATGCTTACTTTTAAACCGTTACCAAGATGATATGCGAGAGAAATTTCTCA 668
QY 686 AAGAGCTTGAAGAGAGATGATGATGCAAGATTCCTTGCACTTTTGAAGGTCACATTC 745
    |||
Db 669 AAGAGCTTGAAGAGAGATGATGATGCAAGATTCCTTGCACTTTTGAAGGTCACATTC 728
QY 746 AGTGTCTGCAAGACCAAGCTTTACTTTGCGATGTTCTTGTTCAGCAATCA 799
    |||
Db 729 AGTGTCTGCAAGACCAAGCTTTACTTTGCGATGTTCTTGTTCAGCAATCA 782

RESULT 12
BU635010
```



```
|||||
Db 703 CGAAGAGATGAAGTGAACATGATGGCTAAGCAAGAAAGCTAAGCTGCATGAGAA 644
QY 567 AATCAAGAGCAAGCACTACAAATGATGAGATGTTATTTAGAAATCTGTCCACAGAGCA 636
Db 643 AATCAAGAGCAAGCACTACAAATGATGAGATGTTATTTAGAAATCTGTCCACAGAGCA 584
QY 627 AGCTCAGATCAATGCTACTTTTAACCGTTTACCAAGATGATCATGCGAGAGAAATTCGA 686
Db 583 AGCTCAGATCAATGCTACTTTTAACCGTTTACCAAGATGATCATGCGAGAGAAATTCGA 524
QY 687 GAGCTTTGAGAGAGAGATGATGATGCAAGTCTTTGCACTTTTGAAGTCAACATTC 746
Db 523 GAGCTTTGAGAGAGAGATGATGATGCAAGTCTTTGCACTTTTGAAGTCAACATTC 464
QY 747 GTCCTTGACAAAGCCAGAGCTTTACTTTGTCGATGTTCTTCGCTTCAGCAATCAACAAAC 806
Db 463 GTCCTTGACAAAGCCAGAGCTTTACTTTGTCGATGTTCTTCGCTTCAGCAATCAACAAAC 404
QY 807 TGAACCTGATGAGAGAGCACTCACTAGAAATTTGACCAACAGAGCTGAGATTGACTTGA 866
Db 403 TGAACCTGATGAGAGAGCACTCACTAGAAATTTGACCAACAGAGCTGAGATTGACTTGA 344
QY 867 GGTCAATTTGAGAGAGATCCAGCGCAGAGAACAGCAATTCCTTTGAGAGAACTTATCA 926
Db 343 GGTCAATTTGAGAGAGATCCAGCGCAGAGAACAGCAATTCCTTTGAGAGAACTTATCA 284
QY 927 AGACACTCGTGAGAGATTTAGAGAGAAAGTCTCGTGCGACTTCGGTGAGAGATGATGCTTA 986
Db 283 AGACACTCGTGAGAGATTTAGAGAGAAAGTCTCGTGCGACTTCGGTGAGAGATGATGCTTA 224
QY 987 ATCAATCAATCTCTCCACAGAGAAACATTAAGTCTCTACAGCTTCTGTTATCTCTATCT 1046
Db 223 ATCAATCAATCTCTCCACAGAGAAACATTAAGTCTCTACAGCTTCTGTTATCTCTATCT 164
QY 1047 CCTCTCTCTCTCTTTGATGATGATTTCAATCGTTGATTTTGTTCACAAAACCTTGT 1106
Db 163 CCTCTCTCTCTCTTTGATGATGATTTCAATCGTTGATTTTGTTCACAAAACCTTGT 104
QY 1107 TTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1166
Db 103 TTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 44
QY 1167 AGTGTGCTCTCTTAAGTTATATATATATATGAAAGCAATGGCT 1209
Db 43 AGTGTGCTCTCTTAAGTTATATATATATATGAAAGCAATGGCT 1

RESULT 14
LOCUS AV785121/c 702 bp mRNA linear EST 28-MAR-2002
DEFINITION AV785121 RAFL6 Arabidopsis thaliana cDNA clone RAFL06-11-P04 3',
            mRNA sequence.
ACCESSION AV785121 GI:19803911
VERSION AV785121.1
KEYWORDS EST.
SOURCE Arabidopsis thaliana (chale crees)
ORGANISM Arabidopsis thaliana
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
          1 (bases 1 to 702)
          Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Saitou,M., Nakajima,M.,
          Oono,Y., Sekurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
          Arakawa,T., Shibata,K., Shinagawa,A., Mutamatsu,M., Hayashizaki,Y.
          and Shinozaki,K.
          Large scale analysis of Arabidopsis full-length cDNA (2002b)
          Unpublished (2002)
          Contact: Motoaki Seki
          Plant Functional Genomics Research Group
          RIKEN Genomic Sciences Center
          3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
          Tel: 81-298-36-4359
```

```
Fax: 81-298-36-9060
Email: mseki@rc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). This clone is in a
modified pBluescript vector as a SclI/XhoI insert. Please visit our
web site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for
further details.

FEATURES
    source
        1..702
            location/Qualifiers
                /organism="Arabidopsis thaliana"
                /mol_type="mRNA"
                /db_xref="taxon:3702"
                /clone="RAFL06-11-P04"
                /dev_stage="plants at various developmental stages from
                germination to mature seeds"
                /lab_host="pH10B"
                /clone_lib="RAFL6"
                /note="Site 1: SclI; Site 2: XhoI; subjected to
                dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
                hr) treatments"

ORIGIN
Query Match 56.9%; Score 700; DB 1; Length 702;
Best Local Similarity 100.0%; Pred. No. 7.8e-172; Indels 0; Gaps 0;
Matches 700; Conservative 0; Mismatches 0;

QY 509 AAGGAGATGAGTGAACATGACATGGCTAAGCAAGAAAGCTAAGCTGCTCATGAGAAA 568
Db 702 AAGGAGATGAGTGAACATGACATGGCTAAGCAAGAAAGCTAAGCTGCTCATGAGAAA 643
QY 569 TCAAGGACAAGACATCAATGATGAGATGTTATTAATCTTGTCCACAAAGCAAG 628
Db 642 TCAAGGACAAGACATCAATGATGAGATGTTATTAATCTTGTCCACAAAGCAAG 583
QY 629 CTCAGATCAATGCTACTTTTAAACGTTACCAAGATGATCATGCGAGAGAAATTCAGA 688
Db 582 CTCAGATCAATGCTACTTTTAAACGTTACCAAGATGATCATGCGAGAGAAATTCAGA 523
QY 689 GTCTTGAGAGAGAGATGATGATGACAAAGTCTTGCACTTTTGAAGTCAACATTCAGT 748
Db 522 GTCTTGAGAGAGAGATGATGATGACAAAGTCTTGCACTTTTGAAGTCAACATTCAGT 463
QY 749 GCTTGACAAGACAGAGCTTTAATTGTGATGTTCTTGTTCAGCAATCAACAAACCTG 808
Db 462 GCTTGACAAGACAGAGCTTTAATTGTGATGTTCTTGTTCAGCAATCAACAAACCTG 403
QY 809 GAATGATGAAGAGACATCACTAGATTTGTACCAAGAGCTGAGATTGCTGGAAG 868
Db 402 GAATGATGAAGAGACATCACTAGATTTGTACCAAGAGCTGAGATTGCTGGAAG 343
QY 869 TCATTGAGAGAGAGTACAGCGCAGAGAACAGATTCCTTTGAGAGAAAGCTATTACCAAG 928
Db 342 TCATTGAGAGAGAGTACAGCGCAGAGAACAGATTCCTTTGAGAGAAAGCTATTACCAAG 283
QY 929 ACACTGTGAGAGATTACGAGAAAGATCTCTGTGCACTTCTCGGTGAAGATGCTTAAT 988
Db 282 ACACTGTGAGAGATTACGAGAAAGATCTCTGTGCACTTCTCGGTGAAGATGCTTAAT 223
QY 989 CAATCAATCTCTCCACAGAGAAACATTAAGTGTCTTACAGCTTCTGTATCTTTATCTCC 1048
Db 222 CAATCAATCTCTCCACAGAGAAACATTAAGTGTCTTACAGCTTCTGTATCTTTATCTCC 163
QY 1049 CTCTGCTCTCTTTGATGATGATTTCAAAATGTTGTTGTTCTTACAAAACCTTGTTT 1108
Db 162 CTCTGCTCTCTTTGATGATGATTTCAAAATGTTGTTGTTCTTACAAAACCTTGTTT 103
QY 1109 GTTTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1168
Db 102 GTTTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 43
QY 1169 TGTGCTCTTTAAGTTATATATATATGAAAGCAATGGCC 1208
Db 42 TGTGCTCTTTAAGTTATATATATATGAAAGCAATGGCC 3
```



RESULT 15  
 BX836971 679 bp mRNA linear EST 11-FEB-2004  
 LOCUS BX836971 Arabidopsis thaliana Adult vegetative tissue Col-0  
 DEFINITION Arabidopsis thaliana cDNA clone GSLT5382F07 5PRIM, mRNA sequence.  
 ACCESSION BX836971  
 VERSION BX836971.1 GI:42531054  
 SOURCE EST.  
 ORGANISM Arabidopsis thaliana (thale cress)  
 Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 679)  
 Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,  
 Menard,M., Cruaud,C., Quelier,F., Scarpelli,C., Schachter,V.,  
 Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
 Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome  
 Annotation  
 Unpublished (2004)  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE  
 Email: seqret@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli  
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for Protein  
 Sequences).  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/EST  
 http://www.genoscope.cns.fr/cgi-bin/gdb/gdb?source=Arabidopsis.  
 Location/Qualifiers  
 1. 679  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLT5382F07"  
 /issue\_type="Adult vegetative tissue"  
 /clone\_lib="Arabidopsis thaliana Adult vegetative tissue  
 Col-0"

ORIGIN

Query Match 54.3%; Score 667.8; DB 5; Length 679;  
 Best Local Similarity 99.0%; Pred. No. 2,1e-163;  
 Matches 672; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 507 CGAAGGAGATGAATGAACATGACATGGCTAAGCAAGAAAGCTGCTCATGAGAA 566  
 DB 1 CGAAGGAGATGAATGAACATGACATGGCTAAGCAAGAAAGCTGCTCATGAGAA 60  
 QY 567 AATCAAGACACAGACATACATGATGAGATGATTAATAGATCTTGTCCAAGAAAGCAA 626  
 DB 61 AATCAAGACACAGACATACATGATGAGATGATTAATAGATCTTGTCCAAGAAAGCAA 120  
 QY 627 AGCTCAGATCAATGCTATCTTTTAAACGTTACCAAGATGATCATGCGAGAAATTCTCAA 686  
 DB 121 AGCTCAGATCAATGCTATCTTTTAAACGTTACCAAGATGATCATGCGAGAAATTCTCAA 180  
 QY 687 GAGCTTGAGAGAGAGATGATGATGACAAAGTTCTTGCACTTTTGAAGTCAACCATTTCA 746  
 DB 181 GAGCTTGAGAGAGAGATGATGATGACAAAGTTCTTGCACTTTTGAAGTCAACCATTTCA 240  
 QY 747 GTGCTTACAAGACAGAGCTTTACTTTGCGATGTTCTTGTGCTTCAAGATCAACAAAC 806  
 |||||

DB 241 GTGCTTACAAGACAGAGCTTTACTTTGCGATGTTCTTGTGCTTCAAGATCAACAAAC 300  
 QY 807 TGGAACTGATGAGAGAGACCTCACTAAGATTGTGACCAAGAGCTGAGATTGACTGAA 866  
 DB 301 TGGAACTGATGAGAGAGACCTCACTAAGATTGTGACCAAGAGCTGAGATTGACTGAA 360  
 QY 867 GGTCAATTGAGAGAGAGATACAGAGGCAAGAAACAGCACTTCTTGTGAGAAAGCTATTACCA 926  
 DB 361 GGTCAATTGAGAGAGAGATACAGAGGCAAGAAACAGCACTTCTTGTGAGAAAGCTATTACCA 420  
 QY 927 AGACACTCGAGAGATTACGAGAAAGATGCTGTCGCACTTCTCGTGAAGATGATGCTTA 986  
 DB 421 AGACACTCGAGAGATTACGAGAAAGATGCTGTCGCACTTCTCGTGAAGATGATGCTTA 480  
 QY 987 ATCAATCAATCTTCCACAGAGAAACATTAAGCTGCTCTACAGCTTGTGTTATCTTATCT 1046  
 DB 481 ATCAATCAATCTTCCACAGAGAAACATTAAGCTGCTCTACAGCTTGTGTTATCTTATCT 540  
 QY 1047 CCTCTCTCTCTCTTTGATGATGATTTCAATCGTTGATTTTGTCTTCAAAAACTTTGT 1106  
 DB 541 CCTCTCTCTCTCTTTGATGATGATTTCAATCGTTGATTTTGTCTTCAAAAACTTTGT 600  
 QY 1107 TTGTTTCTGTTGTGTTTGAAGTTCCTAAATATGCAAAAGAGAGACAGAGAAC 1166  
 DB 601 TTGTTTCTGTTGTGTTTGAAGTTCCTAAATATGCAAAAGAGAGACAGAGAAC 660  
 QY 1167 AGTGTGCTCTTTAAGTTA 1185  
 DB 661 AGTGTGCTCTTTAAGTTA 679  
 |||||

Search completed: August 22, 2005, 10:19:14  
 Job time : 4657 secs



This Page Blank (uspto)



Db 451 TATGCAAGACAAGCATATCATGCCGATACAAAGATCAATGAAAGAGACCTGCTCA 510  
Qy 450 CCAACATACCGGTGACTTTCAGAAAGCTTTGGTTTCTCTGTTAACTCAATACAGTACGA 509  
Db 511 CCAACAACCTGGAGATTGTTGTAAGTTGCTGGTACCTTGGAGGCTCTTACCGTATGA 570  
Qy 510 AGGAGATGAAGTGAACATGACATTTGGCTAAGCAAGAGCTAAGCTGGTCCATGAGAAAT 569  
Db 571 TGAAGATGAGTGAATATGACTTTGGCAAGAGAGGCTAAGATATCTCCACAGAAAGAT 630  
Qy 570 CAAGACACAGCACTACATATGATGAGATGTTATTGAATCTTGTCCAAGAAAGCAAGC 629  
Db 631 CTGAGAGAGGCTTATGCGCATAGAGATCTCATAGGATTTGGCTACTAGAGCAAGC 690  
Qy 630 TCAGATCAATGCTACTTTTAAACGTTAACCAAGATGATCAAGGAGGAAATTTCCAGAG 689  
Db 691 ACGGTGATATGCTACGCTGTAATCACTACAAATATGATTGGAAATGATATCAACAGGA 750  
Qy 690 TCTTGAAGAAAGAGATGATGATGACAAATTCCTTGGCACTTTGAGGTCAACCATTCAGT 749  
Db 751 TTT---GAAATCTGATCCAAAGACGCGTTCCTTACTATCTAGAGCTACAGTAAAGT 807  
Qy 750 CTTGACAAGACCAAGCTTTACTTTTGTGATGTTTCTTGTTCAGCAATCAACAAACTG 809  
Db 808 CCTGACTCGCCCTGAGAGATATTTGAAAAGTTCTTGTCTAGCCATCAATTAAGGAGG 867  
Qy 810 AACTGATGAAGAGCACTCACTGAATTTGACCAACAGAGCTGAGTTGACTTGAAGGT 869  
Db 868 AACGATGAAGGGGCTTGAACAGAGTATGTTCTTCCAGGGCCGAGGTTGATGAAGTT 927  
Qy 870 CATTGAGAGAGTACCAGGCAAGAACAGATTCCTTTGAGAAAAGCTATTACCAAGA 929  
Db 928 TATTAAGTGAAGAGTACAGAGAGGAATAGATCCCTCGATCGAGCATGTGCAAGA 987  
Qy 930 CATCTGGAGATTACAGAAAGTCTCTGCGCACTTCTCGTGAAGATATGCTTATC 989  
Db 988 CACTTACTGGAGACTATGAAAAATGCTTGTGCAATGATGGCCAGCTGAGGCTTGATT 1047  
Qy 990 AATCAATCCTC 1000  
Db 1048 TACAAGTACTC 1058

RESULT 2  
US-09-325-932A-20  
; Sequence 20, Application US/09325932A  
; Patent No. 6451604  
; GENERAL INFORMATION:  
; APPLICANT: Flinn, Barry  
; APPLICANT: Lasham, Annette  
; TITLE OF INVENTION: Compositions affecting programmed cell  
; TITLE OF INVENTION: death and their use in the modification of forestry plant develo  
; FILE REFERENCE: 1022  
; CURRENT APPLICATION NUMBER: US/09/325,932A  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 20  
; LENGTH: 789  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-325-932A-20

Query Match 29.6%; Score 364.2; DB 3; Length 789;  
Best Local Similarity 69.7%; Pred. No. 6,le-97;  
Matches 508; Conservative 0; Mismatches 218; Indels 3; Gaps 1;

Qy 25 AGAAGAAAATGGCGACTCTTAAGTTTCTGATTCCTGCTCTCTCTGATGATGCT 84  
Db 64 AGAAGAGATGGCGACTATCGCGGTCCACCGTCCGTTCCGCTCCGCTGAGATGCC 123  
Qy 85 GAGCAATTGGAACCGCTTTTGAAGATGGGTTACGACGAGACCTTGATCATATC 144

Db 124 GAGCAGCTCCAAAAAGCTTTCCAGAGATGGGGAACGAATGAAGATCTGATCATATCATA 183  
Qy 145 TTGGCTCAGAGAAGTGTCTGAACAGAGAAAGTCAATCAGGCAAGCATACAGAAACCTTC 204  
Db 184 CTGCTCAGAGAAAGCAGACGACGAGAAAGTAAATCCGAAACATATGCCAGACATAT 243  
Qy 205 GCGAAGACCTTCTCAAGACTTTTGAACAAGAGCTCTCTCAAGATTTTCCAGAGACTATC 264  
Db 244 GGGGAAGATCTTCTCAAGACGTTGACAGAGAACTCTCTAGTGACTTTGAGAGATCTG 303  
Qy 265 TTGTTGGAAGCTTTGAACCCGATGACGTTATGCTTTATTTGCTTAATGAGCTACAAA 324  
Db 304 CTTCGTGTGACCTGGATCTCGGAGCGTGTATGCAATCTTGTCCAATGAATCAAG 363  
Qy 325 AATAGACTCTTAAGCAACCAAGTTCTTATGAGTTGCTTGAACAAGACATCAACGCG 384  
Db 364 AATAGACTCTTAAGCAACCTGAGTTCTCAAGAAATTTGCAACGAGCTTCAATGAG 423  
Qy 385 CTGCTTCAAGCTAGGCAAGCTTACATGCTCGCTCAAGAACTCTTTGAAGAGACGTT 444  
Db 424 TTAATCATGTGAGGAGGCTATATGCTCTGTTAAGAAATCTTGAAGAGACATC 483  
Qy 445 GCTCACACACTACCGGTGACTTCAAGAAAGCTTTTGTGTTTCTTGTTAATCTCATACG 504  
Db 484 GATATCACTACATGAGGATTTCCGCAAGCTGCTTGTCTCTGCAAGTACCTTTCG 543  
Qy 505 TCGAAGAGATGAAGTGAACATGACATTTGGCTAAGCAAGAGCTAAGCTGATGATG 564  
Db 544 TATGAGGGGCTGAGTGAACATGATTTGGAGATCAGAGGCTAAGTATCTATGAG 603  
Qy 565 AAATCAAGACAGCACTACATGATGAGATGTTATGAAATCTTGTCCCAAGAAC 624  
Db 604 AAGATCAAGAGAGCTTACATCATATGATGAGCTCATGAAATGTTATCAAGAGT 663  
Qy 625 AAAGCTCAGATCAATGCTTCTTAAACGTTACCAAGATGATACGAGAAATCTTC 684  
Db 664 AAAGCTCAGCTTAATGCAACCTCAATTAATCAACAAAGATTTGGAAATGCAATCAAC 723  
Qy 685 AAGAGCTTGAAGAGAGATGATGATCAAGTTCCTTCACTTTGAGGTCAACCAT 744  
Db 724 AAGATCT---GAGGCTGATCCAAATGATGATTTGAAACCTGAGATCAGCAAT 780  
Qy 745 CAGTCTTG 753  
Db 781 AAGTCTTG 789

RESULT 3  
US-09-325-932A-21  
; Sequence 21, Application US/09325932A  
; Patent No. 6451604  
; GENERAL INFORMATION:  
; APPLICANT: Flinn, Barry  
; APPLICANT: Lasham, Annette  
; TITLE OF INVENTION: Compositions affecting programmed cell  
; TITLE OF INVENTION: death and their use in the modification of forestry plant develo  
; FILE REFERENCE: 1022  
; CURRENT APPLICATION NUMBER: US/09/325,932A  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 21  
; LENGTH: 704  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-325-932A-21

Query Match 16.9%; Score 208.4; DB 3; Length 704;  
Best Local Similarity 60.4%; Pred. No. 5,le-51;  
Matches 344; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

Qy 6 CGTCCGAACACTAAAGTGAAGAAAATGGCGACTCTTAAGGTTTCTGATTCGTTCC 65

Db 120 CGTGGAGTAGTGTGATTTGTCGAGAAATGTCAACATTAATAGTCCAGTTCCAAATACC 179  
QY 66 TGTCTCTTGTGATGATCTGAGCAATTGAGAACCGCTTTTGAAGATGGGTATGAGAACGA 125  
Db 180 GACCCCATCTGAAACCTCTGAACCCGAGAGAAAGCTTTGAAAGGGGGGACAAATGA 239  
QY 126 GGAATTATATATCAATCTTGGCTCAAGAAAGTGTGAACAGAGAAAGCATCAGGCA 185  
Db 240 GAAGTCAATCATCAAAATATTAGACATTAAGACTGACACACAGCAAGTAATCCGTCA 299  
QY 186 AGCATACCAAGAAACCTACGCGGAGACCTTCTCAAGACTCTTGAACAAGAGCTCTTAA 245  
Db 300 AAGTTATTTCACTGATACGAAAGAGATCTTGAACGATTAAGATCTGAGCTTTCAAG 359  
QY 246 CGATTTGAGAGAGCTATCTTGTGTGACTCTTGAACCGGTGAGGTGATGCTTTATT 305  
Db 360 TGACTTTGAGAAAGAGCTATCTTGTGTGATCTGATACAGCTGAAAGTGTGAGGTAT 419  
QY 306 GGTATGAGAGCTACAAAAGATGAGCTTCAAGCAACCAAGTTCTTATGAAAGTTGCTTG 365  
Db 420 ATTCATGTGCAATTAAGATGAGATGCAAGAAATATATCGCTTTAGAAATTTCCAG 479  
QY 366 CACAAGACATCAACGAGCTGCTTCAAGCTAGGCAAGCTTACCAGCTGCTACAAAGAA 425  
Db 480 TGTCTGATCTTGGCTGAATGATGATGAGGCAAGCATATCATATTCGTACAAA 539  
QY 426 GTCTCTTGAAGAGAGCTTGTCTACCACTACCGGTGACTTCAAAAAGCTTTGGTTTC 485  
Db 540 GTCCCTCGAAAGACGTGCTGCAATACAAAGTGAACCTCCGTAAAGTTGCTGATAGC 599  
QY 486 TTTTGTACCTCATACAGATGACGATGAGATGAAAGTGAATACATATTTGCTTAAGCAGA 545  
Db 600 ACTTGTAGTTCAATGATGATGATGAGGTCCGAAAGTGAATGATTTGGCAAGTTATGA 659  
QY 546 AGCTAAGCTGTGCATGAGAAATCAAGGA 575  
Db 660 AGCAAGAAAGCTAAGTGAATCTATTAACGA 689

RESULT 4  
US-09-325-932A-18  
; Sequence 18, Application US/09325932A  
; Patent No. 6451604  
; GENERAL INFORMATION:  
; APPLICANT: Film, Barry  
; APPLICANT: Lasham, Annette  
; TITLE OF INVENTION: Compositions affecting programmed cell  
; FILE OF INVENTION: death and their use in the modification of forestry plant develo  
; FILE REFERENCE: 1022  
; CURRENT APPLICATION NUMBER: US/09/325,932A  
; NUMBER OF SEQ ID NOS: 206  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 484  
; TYPE: DNA  
; ORGANISM: Pinus radiata  
US-09-325-932A-18

Query Match 15.1%; Score 186.2; DB 3; Length 484;  
Best Local Similarity 74.9%; Pred. No. 1.4e-44;  
Matches 233; Conservative 0; Mismatches 78; Indels 0; Gaps 0;  
QY 254 AGAGACTATCTTGTGTGACTCTTGAACCGGTGAGCTGATGCTTTATTTGGCTAATG 313  
Db 174 AGAGCTGTGTGCTTCTTGTGCTGATCGGTGAGAGTGAATGCTTGGCGAATG 233  
QY 314 AAGCTTCAAAAAGATGATCTTCAAGCAACCAAGTTCTTATGAGATGCTTCAAGAA 373  
Db 234 AAGCGAGAAAGATGATCTTCAAGCAACCAAGTTCTTCAAGAAATGCTTCAAGAGAT 293  
QY 374 CATCAAGAGCTTCAAGCTGAGCAAGCTTACATGCTGCTCAAGAAATGCTTTG 433

Db 294 CTCACAGACGTTGCTCATGTCGCAAGAACAGCATATGCTGTGATCAAGAAAGTCCGTGG 353  
QY 434 AAGAGAGCTTGTCTCACACACATTAACGGTGACTTCAAGAAAGCTTTGGTTTCTTGTGA 493  
Db 354 AAGAGAGCTGCTGTACACACATGAGATTTGTGAAGTTGCTGTACCTTGTGTGA 413  
QY 494 CTTCAATCAGGTGCAAGAGAGTGAAGTGAACATGATTTGGCTTAAGCAAGAGCTTAAGC 553  
Db 414 GCTTCTTCAATTAATGATGAGATGAGTGAATATGATTTGGCAAAAGAGGCTTAAGA 473  
QY 554 TGTTCATGAG 564  
Db 474 TACTCCAGAG 484

RESULT 5  
US-09-949-016-2081  
; Sequence 2081, Application US/09949016  
; Patent No. 6812319  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2081  
; LENGTH: 1332  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-2081

Query Match 12.1%; Score 149.2; DB 4; Length 1332;  
Best Local Similarity 49.0%; Pred. No. 2.1e-33;  
Matches 456; Conservative 0; Mismatches 468; Indels 6; Gaps 2;  
QY 54 TGATTCGTCTGCTCTCTTGTGATGATGTCGATGATGAGAACCGCTTTGAAAGATG 113  
Db 86 TTATTCAGACTTATGACCATGATGATGATGATGATGATGATGATGATGATGATGAT 145  
QY 114 GGTATGAAACGAGACTTATGATCAATCTTGGCTCAAGAAAGTCTGAAACAGAGAA 173  
Db 146 TGGAACTGATGAGAAATGCTATGATGATGATGATGATGATGATGATGATGATGAT 205  
QY 174 AGTATCAAGGCAAGCAATACAGAAACCTACGCGGAGAGCTTCTCAAGATCTTGACAA 233  
Db 206 GCTGATTTGTAAGAAATATCAAGCAATATGAGAAAGAGCTGAAAGATGATGAGAG 265  
QY 234 GAGCTCTTAAAGATTTGAGAGAGCTATCTTGTGTGAGACTTTGAACCGGTGAGAG 293  
Db 266 TGATCTCTGTGCACTTTGAGAGATCTATGATGATGATGATGATGATGATGATGAT 325  
QY 294 TGATCTTATTTGGCTAATGAAAGCTTCAAAAGATGAGATTTCAAGCAACCAAGTTAT 353  
Db 326 TGAATGAAAGAGCTAAGAAATCCATGAAAGGCGCGGAGAAACGAAGATGCTTGAAT 385  
QY 354 GGAAGTGTCTGCAAGAGCATCAACGAGCTGCTTCAAGGCAAGCTTACCAATGCAATG 413  
Db 386 TGAATCTTAACTACAGAGCAAGAGCAAGAAATGAAAGATATCTTCAAGCTTATTAAC 445  
QY 414 TGGCTAAGAAAGTCTTTGAAGAGAGCTTGTCTCAACCACTACCGGTGATCTTGAAGAA 473  
Db 446 AGTATCAAGAAAGAGTCTTGAAGAGATGATGATGATGATGATGATGATGATGATGAT 505  
QY 474 GCTTTGGTTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 533

```

Db      506 AGCTGTTGACCTTTGGCAGATGGCAGAGAGATGAAAGTCTGAAAGTGATGACATCT
QY      534 GGTAGCAAGAGCTAAGTGTGCTCATGAGAAATCAAGACA---AGACTACAAATGA 590
Db      566 GGGCAAAACAGAGCCAGATTTCTATAAAGCTGTGAGAACAGATGGGGCAGATGA 625
QY      591 TGAGATGTTATTAGATCTTGTCCACAAAGAACAAAGCTCAGATCAATGCTACTTTAA 650
Db      626 AGCAAAATTCATGAGATCTCTGTGTTAAGAGCTTTCTCAATTAATAACTAATTTGA 685
QY      651 CCGTTACCAAGATGATCAGCGCAGAAATTTCTCAAGAGCTTTGAGAAAGAGATGATGA 710
Db      686 TGAATACAGAAATATCAGCCAAAGACATTTGACAGCAT---AAAAGCAATTAATC 742
QY      711 TGCAAGTTCCTTGACCTTTGAGTCAACCATTCAGTCTTTCACAAAGCCAGAGCTTA 770
Db      743 TGGGCAATTTTGAAGACTTACTGTGGCCATAGTTAATGTGTAGAGAACACCGCGCCTT 802
QY      771 CTTGTGCAATGTTCTTCTGTTCAAGATCAACAAACTGAACTGATGAGAGCACTCAC 830
Db      803 TTTAGCCGAAAGCTGATCAGACCTTTGAAGGTATTGAACTGATGATTTACTTTGA 862
QY      831 TAGAATTTGACCAAGAGCTGAGATTGACTTGAAGTCAATTGAGAGAGTACCAAGC 890
Db      863 CCGAATTAATGTGTCCAGATCAGAAATTTGACCTTTTGACATTCGAACAGAGTTCAAGA 922
QY      891 CAGGAACAGATTCCTTTGAGAAAGCTATTACCAAGACACTCGTGGAGATTACAGAA 950
Db      923 GCATTTGTGCTATTTCCCTATATTACGCAATTAATCGGATTACTTGGAGACTATGAAT 982
QY      951 GATGCTCGTGCACCTTCTCGTGAAGATGA 980
Db      983 CACACTTTAAATCTGTGTGAGATGA 1012

```

```

RESULT 6
US-09-023-655-1149
; Sequence 1149, Application US/09023655
; Patent No. 6607879
;
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555

```

```

; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1149:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1339 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9186967
;
US-09-023-655-1149
Query Match      12.1%; Score 149.2; DB 4; Length 1339;
Best Local Similarity 49.0%; Pred. No. 2.1e-33;
Matches 456; Conservative 0; Mismatches 468; Indels 6; Gaps 2;
QY      54 TGATTTCTTCTGCTCTCTTCTGATGATGCTGAGCAATTTGAGAACCGCTTTGAAAGATG 113
Db      88 TTATCCAGACTTTTAAAGCCATCAGTGAATGCTGAAGCTATTCAAAAGCAATCAGAGAAAT 147
QY      114 GGGTACGAAGAGACTTGATCATATCAATCTTGGTCAAGAGTGTGAACAGAGAA 173
Db      148 TGAACTGATGAGAAATGCTCATCAGCATTTCTGACTGAGAGCTCAATGCAAGCGGA 207
QY      174 AGTCATCAGGCAAGCATACCAAGAACTTACGGCAAGACCTTTCAGAGCTTTGACAA 233
Db      208 GCTGATTTGTAAGAAATATCAGAGCATATGAAAGAGCTGAAGAAATGATGACTTGAAGGG 267
QY      234 GGAAGCTCTTAAGATTTGAGAGACTTCTTGTGTGAGACTTTGAAACCGGTGAGGG 293
Db      268 TGATCTCTCTGACACTTTGAGCATCTCATGTGTGCCCTTAGTACTTCCACGAGCTT 327
QY      294 TGATGCTTATTTGGCTAATGAGCTAACAAAAGATGAGACTTCAAGCAACCAAGTCTTAT 353
Db      328 TGATGCAAGCAGCTTAAGAAATCCATGAAAGGCGCGGAACAAAGAGATGCTTGAAT 387
QY      354 GGAAGTTCCTTCAAGAGCATCAACGAGCTGCTTCAAGCTAAGCAAGCTTACCATGC 413
Db      388 TGAAATCTTAATCAACAGCAAGCAAGCAATGAAGATATCTCTCAAGCTTATATATC 447
QY      414 TGGCTTCAAGAGCTCTTTGAAGAGACTTCTCAACCATCAACCGGTGATTTCAAGAA 473
Db      448 AGTATACAAAGAGAGCTTTGAGAGTCAATTTGTTCCGAACATCTGTGATCTTCGGAA 507
QY      474 GCTTTGGTTCTCTTGTATACCTCATACAGTACGAGAGAGATGAGATGATGATCAT 533
Db      508 AGCTCTGTGACTTTGGCAGATGGCAGAGAGATGAAAGTCTGAAAGTGAATGAGATCT 567
QY      534 GGTACGCAAGAGCTTAAGCTGTGCTCATGAGAAATCAAGACA---AGACTTACATGA 590
Db      568 GGGCAAAACAGATGCCAGATTTCTATTAAGCTGTGTGAGAACAGTGGGGCAGCATGA 627
QY      591 TGAGATGTTATTAGATCTTGTCCACAAAGAGCAAGCTCAGATCAATGCTACTTTTA 650
Db      628 AGCAAAATTCAGTGAATCTCTGTGTTAAGAGCTTTCTCAATTAATACTAATTTGA 687
QY      651 CCGTTACCAAGATGATCAGCGCAGAAATTTCTCAAGAGCTTTGAGAAAGAGATGATGA 710
Db      688 TGAATACAGAAATATCAGCCAAAGACATTTGAGACAGCAT---AAAAGCAATTAATC 744
QY      711 TGCAAGTTCCTTGACCTTTGAGTCAACCATTCAGTCTTGAACAGACCAAGAGCTTTA 770
Db      745 TGGGCAATTTTGAAGACTTACTGTGGCCATAGTTAATGTGTGAGAAACACCGCGCCTT 804
QY      771 CTTGTGCAATGTTCTTCTGCTGATCAACAAACTGAACTGATGAGAGAGCACTCAC 830
Db      805 TTTAGCCGAAAGCTGATCAGACCTTTGAAGGTATGGAAGCTGATGATTTAATCTTTGA 864
QY      831 TAGAATTTGACCAAGAGCTGAGTGAATTTGAGAGTCAATTTGAGAGAGTACCAAGC 890
Db      865 CCGAATTAATGTGTCCAGATCAGAAATTTGACCTTTTGACATTTCAAGAGATTTCAAGAA 924
QY      891 CAGGAACAGATTCCTTTGAGAAAGCTATTACAAAGCACTGTGAGATTTACAGAA 950

```



```
/ Patent No. 6607879
/ GENERAL INFORMATION:
/ APPLICANT: Cocks, Benjamin G.
/ APPLICANT: Susan G. Stuart
/ APPLICANT: Jeffrey J. Seilhamer
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
/ TITLE OF INVENTION: EXPRESSION
/ NUMBER OF SEQUENCES: 1508
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/023,655
/ FILING DATE: HEREWITH
/ CLASSIFICATION:
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0001 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 1152:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1362 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GENBANK
/ CLONE: g187137
/ US-09-023-655-1152

Query Match      8.4%; Score 103.8; DB 4; Length 1362;
Best Local Similarity 46.7%; Pred. No. 5.3e-20;
Matches 435; Conservative 0; Mismatches 487; Indels 9; Gaps 3;
```

```
QY 416 GCTACAGAAGTCTCTTGAAGAGAGCGTTGCTCACCACACTACCGGTACTTCAGAAAGC 475
    |||||
DB 498 TGTACAGACTGATCTGAGAGAGACATTTATTTGGACACATCTGTGATCTTCGGACAC 557
QY 476 TTTGGTTTCTCTTGTACCTCATACAG---GTACGAAGAGATGAAGTGAACATGACAT 532
    |||||
DB 558 TGATGGTTCCCTGGCAAAAGGTTAGAAAGCAGAGATGGCTCTGTCAATGATTTATGAAC 617
QY 533 TGGCTAAGCAAGACTAAGCTGTGTCATGA---GAAATCAAGCAAGCACTACAAATG 589
    |||||
DB 618 TGATTGACCAAGATCTCGGATCTCTATGAGCTGTGAGTGAAGGAAAGAACTGATG 677
QY 590 ATGAGGATGTATTTGAATCTTGTCCACAAGAGCAAGCTGATCAATGCTTACTTTTA 649
    |||||
DB 678 TTCCCAAGTGATCAGCATCATGACCGAGCGAGGTGCCCATCCAGAAAGATATTTG 737
QY 650 ACCGTTACCAAGATGATCATGCGAGAAATTTCTCAAGATCTTGAAGAAAGATGATG 709
    |||||
DB 738 ATAGGTACAAAGATTACAGCCCTTATGACATGTTGAAAAGCATCAGAAAAGGTTAAAG 797
QY 710 ATGACAAGTCTTGTGACATTTTGAAGTCAACATTCAGTCTTGAACAAGACAGACTTT 769
    |||||
DB 798 GAGACCTGGAATAATC---TTTCTGAACTGTGATGATTCAGAACAAAGCCCTGT 854
QY 770 ACTTGTGATGTTCTGCTTCAAGCAATCAACAAACTGSAATGATGAAGAGCACTCA 829
    |||||
DB 855 ATTGTTGCTGATGGCTGTATGATCTCATGAAAGGCAAGGAGATGAAGTCTCTGA 914
QY 830 CTAGAATTGTGACCAAGAGCTGATGATGATCTTGAAGTCAATTTGAAGAGATACAC 889
    |||||
DB 915 TCAGATATCATGCTCTCCGACGTGAAGTGAACATGTTAAATTAAGTCTGAATTCAGA 974
QY 890 GCAGAAACAGCATTTCTTTGAGAAAGCATATACCAAGACATCTGTGAGATTACGAGA 949
    |||||
DB 975 GAAAGTACGGCAAGTCCCTGTATCTATATTCACCAAGACACTAAGGCGCACTACAGA 1034
QY 950 AGATGCTGTCGCACTTCTCGGTGAAGATGA 980
    |||||
DB 1035 AAGCGCTGTATCACTGTGTGTGAGATGA 1065

RESULT 10
US-09-814-915A-101
/ Sequence 101, Application US/09814915A
/ GENERAL INFORMATION:
/ APPLICANT: Horwitz, Kathryn
/ APPLICANT: Richer, Jennifer
/ TITLE OF INVENTION: Progesterone Receptor--Regulated Gene Expression and Methods Relat
/ TITLE OF INVENTION: Threoto
/ FILE REFERENCE: 2848-39
/ CURRENT APPLICATION NUMBER: US/09/814,915A
/ PRIOR FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 60/214,870
/ NUMBER OF SEQ ID NOS: 108
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 101
/ LENGTH: 1362
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-814-915A-101

Query Match      8.4%; Score 103.8; DB 4; Length 1362;
Best Local Similarity 46.7%; Pred. No. 5.3e-20;
Matches 435; Conservative 0; Mismatches 487; Indels 9; Gaps 3;
```



Db 198 GTGTGATGAGTACCATTTCTCAACATTTTGAACCAACCGAGCAATGCAAGAGACAG 257  
Qy 176 TCATCAGCAAGATACCAAGCAAACTTACCGGGAAGACTTTCTCAAGACTTTGACAAAG 235  
Db 258 ATATTGCTTGGCTTCAAGAGAGACCAAAAAGAACTTGATCAGTCACTGAAGTCAAG 317  
Qy 236 AGCTCTTAAGATTTTGAAGAGACTTATCTTGTGTGATGACTTTGAAACCGGTGAAGCTG 295  
Db 318 CTTATCTGGCCACCTGAGACGGTGTATTTGGGCTTATTTGAAGACCTGCTCAGTATG 377  
Qy 296 ATGCTTATTTGGCTATATGAAGCTCAAAAAGATGATGACTTCAAGCAACCAAGTTCTATGG 355  
Db 378 AGCTTTCTGAGCTTAAAGCTTCCATGAAAGGAGCTGGGAACCGAGAGACTTCTTCATGG 437  
Qy 356 AAGTTGCTTGCAAGAGCATCAACGAGCTGCTTCAAGCTGAGCAAGCTTACCTAGTCTC 415  
Db 438 AGATCATCTGCTCAGAACCAACGAGAGCTGAGAAATTTACAGAGTCTACAGAGAA 497  
Qy 416 GCTTCAAGAAATCTTGAAGAGAGCTTGTCTACACACTACCTGCTGACTTCAAGAAAG 475  
Db 498 TGTACAAAGACTGATCTGAGAAAGACATTTATTTGACACATCTGTGACTTCCGCAAG 557  
Qy 476 TTTTGGTTTCTCTTGTATACCTCATACAG--GTACGAAGGATGAAGTGAACATGACAT 532  
Db 558 TGAATGTTGCTTGGCTGCAAGAGGTGAAGAGCAGAGATGGCTGTCTATGATTTATGAC 617  
Qy 533 TGGCTAAGCAAGAGCTTAAAGCTGTGCTCATGTA--GAAATCAAGAGCAAGCACTACATG 589  
Db 618 TGAATGCAAGAGATGCTGCGGATCTTATGACGCTGAGTGAAGAGAAAGAACTGATG 677  
Qy 590 ATGAGATGTTATTAAGATCTTGTCCACAGAGCAAAAGCTCAGATCAATGCTACTTTTA 649  
Db 678 TTCCCAAGTGAATCAGATCATGACCGAGCGAGCTGCCCCACCTCAGAAAGTATTTG 737  
Qy 650 ACCGTTCACAGATGATCATGAGGGAATTTCAAGAGCTTGAAGAGAGAGATGATG 709  
Db 738 ATAGATCAAGAGATTAACAGCCCTTATGATGATTTGAAGAGATTAAGAAAGGTTAAG 797  
Qy 710 ATGACAAATTCCTTGCATTTTGAAGTCAACATTCAGTGTGTAACAAGCAAGAGCTTT 769  
Db 798 GAGACCTGGAAGAAATGC---TTTCTGAACCTGTGATGATTCAGAAACAGCCCTGT 854  
Qy 770 ACTTGTGATGTTCTTGTTCAGCAATCAACAAACTGAACTGATGAAGAGAGACTCA 829  
Db 855 ATTTTGCATCGGCTGATGATCTCAATGAAGGCAAGGGAACGAGATTAAGTCTCTGA 914  
Qy 830 CTAAATTTGTAACCAAGAGCTGAGATGACTTGAAGGTCAATTGAGAGAGATACAGC 889  
Db 915 TCAGAAATCATGCTTCCCGAGTAGAGTGAATGAAATTAAGTCTGAATTTCAAGA 974  
Qy 890 GCAGAAACAGATTCCTTTGGAAGAAAGCTTATTAACCAAGACACTGAGAGATTAACAGA 949  
Db 975 GAAAGTACGGAATCCCTGTACTTATTAATCCAGCAAGACATTAAGGGGCACTACAGA 1034  
Qy 950 AGATGCTGTGCACTTCTCGTGAAGATGA 980  
Db 1035 AAGCGCTGCTGTACTGTGTGTGAGATGA 1065

RESULT 11  
US-09-949-016-2431  
; Sequence 2431, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20

Qy 56 ATTGTTCTGCTGCTCTTCTGATGATGCTGAGCAATTTGAACCCGTTTGAAGATGAG 115  
Db 205 ATACTTAATTTGATGCTGAGCGGATGCTTTGAACATTTGAACAGCATCAAGACCAAG 264  
Qy 116 GTACGAACGAGACTTGAATCATATCAATCTTGCTCAAGAGTGTGAACAGAGAAAG 175  
Db 265 GTGTGATGAGTCAACATTTGTCAACATTTTGAACCAACGAGCAATGACAGAGACAG 324  
Qy 176 TCATCAGCAAGATACCAAGAACTAGCGGGAAGACCTTCTCAAGACTCTTGAAGAG 235  
Db 325 ATATTGCTTGGCTTCAAGAGAGAGCAAAAAGAACTTGCATCAGACTGAAGTCAAG 384  
Qy 236 AGCTCTTAAGCAATTTGAGAGAGCTATCTTGTGTGACTCTTGAACCCGCTGAGCGTG 295  
Db 385 CTTATCTGCGCACCTGAGAGCGGTATTTGGGCTTATTTGAAGACACTGCTCAGATG 444  
Qy 296 ATGCTTATTTGCTTAATGAAGCTTCAAAAAGATGACTTCAAGCAACCAAGTTCTTATG 355  
Db 445 AGCTTCTGAGCTTAAAGCTTCAATGAAGGCTGGAACCGAGAGACTCTCTCATG 504  
Qy 356 AAGTGTCTGCAAGAGCATCAAGCAGCTGCTTCAAGCTGAGCAAGCTTACATGCTC 415  
Db 505 AGATCATCTGCTCAAGAACCAACAGAGCTGAGAGAAATTTACAGAGTCTCAAGAGAA 564  
Qy 416 GCTACAAAGATCTCTTGAAGAGAGCTTGTCTACACACTACCGTGTGACTTCAAGAAAG 475  
Db 565 TGTACAAAGTATCTGAGAGAGAGACATTTTTCGACACATCTGATGCTTCCGAGC 624  
Qy 476 TTTTGGTTTCTTGTATTAACCTCATACAG--GTACGAAGAGATGAAGTGAACATGACAT 532  
Db 625 TGAATGTTGCTTGGCAAGGAGTGAAGAGAGATGAGTCTGTCTGATTTATGAAC 684  
Qy 533 TGGCTAAGCAAGAGCTTAAAGCTGTGCTCATGTA--GAAATCAAGAGCAAGCACTACAAAG 589  
Db 685 TGAATGACCAAGATGCTCGGATCTCTATGACGCTGAGTGAAGAGAGAAAGAACTGATG 744  
Qy 590 ATGAGATGTTATTAAGATCTTGTCAAGAAAGCAAGCTGAGATCAATGCTACTTTTA 649  
Db 745 TTCCCAAGTGAATCAGATCATGATACCGAGGAGCTGCTCCCACTTCCGAAAGTATTTG 804  
Qy 650 ACCGTTCACAGATGATCATGAGCAAAATTTCAAGAGTCTTGAAGAGAGATGATG 709  
Db 805 ATAGATCAAGAGTTAACAGCCCTTATGATGATTTGAAGAGCATCAGAGAAAGGTTAAG 864  
Qy 710 ATGACAAATTCCTTGAATTTGAGGTCAACATTTAGTGTGCTTGAACCAAGAGCTTT 769  
Db 865 GAGACCTGGAAGAAATGC---TTTCTGAAACCTGCTGAGTCAATTCAGAAACAGCCCTGT 921  
Qy 770 ACTTGTGATGTTCTTGTGAGCAATCAACAAACTGAACTGATGAAGAGAGACTCA 829  
Db 922 ATTTGCTATGCTGCTGTATGACTCATTAAGGAGGAGGAGCGAGATTAAGTCTCTTA 981  
Qy 830 CTAAATTTGTAACCAAGAGCTGATGATGACTTGAAGGTCAATTTGAGAGAGATCAAGC 889  
Db 982 TCAGAAATCATGCTTCCCGAGTGAAGTGAATGTTGAATTAAGTCTGAATTTCAAGA 1041  
Qy 890 GCAGGAACAGATTCCTTTGAGAAAGCTATTAACCAAGACACTGCTGAGATTTACAGA 949

PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 2431  
; LENGTH: 1614  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-2431

Query Match 84%; Score 103.8; DB 4; Length 1614;  
Best Local Similarity 46.7%; Pred. No. 5.9e-20;  
Matches 435; Conservative 0; Mismatches 487; Indels 9; Gaps 3;

```
Db      1042 GAAAGTACGGCAAGTCCCTCTACTATATATCCAGCAAGACACTAAGGGGACTACAG 1101
Qy      950 AGATGCTGTCGCACTTCGCTGAAGATGA 980
Db      1102 AAGCGCTGCTGCTGCTGTGTGTGAGATGA 1132

RESULT 12
US-09-919-039-393
; Sequence 393, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaseer, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PP-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 393
; LENGTH: 1660
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 132602.6
US-09-919-039-393
```

```
Query Match      8.4%; Score 102.8; DB 4; Length 1660;
Best Local Similarity 46.7%; Pred. No. 1.2e-19;
Matches 435; Conservative 0; Mismatches 487; Indels 10; Gaps 3;
```

```
Qy      56 ATTCTGTTCTCTCTCTCTCTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGG 115
Db      238 ATACTAACTTTGATGCTGAGCGGGATGCTTTGAACATTGAAACAGCATCAAGACCAAG 297
Qy      116 GTACGAACGAGGACTTGATGATATCAATCTTGCTCAGCAAGATGCTGAAAGAGAAAG 175
Db      298 GTGTGATGAGGCTCACCATTGTCAACATTTTGACCAACCGCAAGATGACAGAGACAG 357
Qy      176 TCATCAGGCAAGATACCAAGAACTTACGCGCAAGACCTTTCAGAGACTTTGACAAAG 235
Db      358 ATATTGCTTCGCTCTACAGAGAGAGACCAAAAGAACTTGATGACAGACTGAATCAG 417
Qy      236 AGCTCTCTTAAGATTTGAGAGAGCTATCTTTGTGGACTTTGAACCGGTGAGCGTG 295
Db      418 CCTTATCTGGCCACCTGAGAGACCTGATTTTGGGCTTATGAGAGACACCTGCTCAGTATG 477
Qy      296 ATGCTTATTGSGTAAATGAGCTACAAAAGATGACTTCAAGCAACCAAGTTCTTATG 355
Db      478 AGCCTTCTGAGTAAAGCTTCCATGAAAGGCGCTGGAAACGACGAGAGCTCTCTCATTTG 537
Qy      356 AAGTCTCTGACAAAGACATCAACGACGCTGCTTACGCTAGGCAAGCTTACATGCTC 415
Db      538 AGATCATCTGCTCCAGAACCAACGAGAGCTGACGAAATTAACAGATCTACAGAGAA 597
Qy      416 GCTACAGAAATCTTTGAAGAGAGCTTGTCACCAACTACCGGTGACTTCAAGAAC 475
Db      598 TGTACAAAGACTGATGAGAGAGACATTAATTCGAGACATCTGATCTTCCGCAAG 657
Qy      476 TTTTGTCTCTCTTTTACTCTCATACAG---TACGAAGAGATGAATGAATGATGA 531
Db      658 TGTATGTGCGCTTGGCAAGGGGTAGAAAGAGAGAGATGGCTCTCTCATTTGATTAGAA 717
Qy      532 TTGGCTAAGCAAGATGATGATGCTCAAGA---GAAATCAAGGACAAGCACTACAT 588
Db      718 CTGATTTGACCAAGATGCTGGGATCTCTATGACGCTGGAGTGAAGGAAAGAACTGAT 777
Qy      589 GATGAGATGTTATTAAGATCTTTGTCACAAAGAGCAAAAGCTCAATCAATGCTACTTTT 648
```

```
Db      778 GTTCCCAAGTGATGAGATCATGACCGAGCGGACGTCGCCCACTCCAGAAAGATATT 837
Qy      649 AACCGTTACCAAGATGATGATGATGAGGAAATTTCCAGAGTCTTGAGGAAGAGATGAT 708
Db      838 GATGATGACAGAGATTAACGCCCTTATGATCATGTTGAGAAAGATACAGAAAGGTTAAA 897
Qy      709 GATGACAAATTCCTTGACCTTTTGAGGTCAACCATTCAGTGTGTTGACAAAGACAGCTT 768
Db      898 GGAGACCTGGAANAAGC---TTTCGAAACCTGGTTAGTGTGATTCAGAAACAGCCCTG 954
Qy      769 TACTTTGTGATGTTCTTCTGTTACAGCAATCAAAACCTGGAATGATGAGAGACATC 828
Db      955 TATTTTGTGATGCTGCTGATGATCATGCAAGGCGCAAGGAGCCGAGATTAAGTCTTG 1014
Qy      829 ACTGAAATTTGACCAACAAGAGCTGAGATGATGCTTGAAGGTCATTTGAGAGATGACAG 888
Db      1015 ATCAGATATGATGTTCCCGCAGTGAAGTGCATGTTGAAATTAAGTCTGAATTTCAAG 1074
Qy      889 CCGAGAACAGCATTCCTTTTGAGAAAGCTATTACCAAGACACTGCTGAGATTAAGAG 948
Db      1075 AAAAAAGTACGCAAGATCCCTGTAATTAATTCAGCAAGACACTAAGGGGACTACAG 1134
Qy      949 AAGATGCTGCTGCACTTCTCGTGAAGATGA 980
Db      1135 AAGCGCTGCTGTACTCTGTGTGTGAGATGA 1166
```

```
RESULT 13
US-09-324-096A-1
; Sequence 1, Application US/09324096A
; Patent No. 6323313
; GENERAL INFORMATION:
; APPLICANT: Tait, Jonathan
; APPLICANT: Brown, David
; TITLE OF INVENTION: ANNEXIN DERIVATIVE WITH ENDOGENOUS CHELATION SITES
; FILE REFERENCE: USFW-1-13841
; CURRENT APPLICATION NUMBER: US/09/324,096A
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent version 3.0
; SEQ ID NO 1
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(981)
US-09-324-096A-1
```

```
Query Match      8.3%; Score 101.6; DB 3; Length 981;
Best Local Similarity 46.0%; Pred. No. 1.9e-19;
Matches 418; Conservative 0; Mismatches 484; Indels 6; Gaps 2;
```

```
Qy      76 GATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGGTAGACGAGAGACTTGAATC 135
Db      76 GGTATGAGAAATCTTCGGAAGCTATGAAGGCTTGGCAAGATGAGAGAGATC 135
Qy      136 ATATCAATCTTGCTCAAGAAAGTCTGAACAGAGAAAGTCAATCAGGCAAGCATACAC 195
Db      136 CTGACTCTGTGACATCCGAAATGCTCAGGCGCCAGAAATCTTGACAGCTTTTAA 195
Qy      196 GAAACCTACGCGGAGAGACCTTCAAGACTCTTGCAAGAGAGCTCTCAAGATTTGAG 255
Db      196 ACTCTGTTTGGAGGAGATCTTCTGATGATCACTGAATGAGAACTTACTGAAAAATTTGAA 255
Qy      256 AGAGCTATCTTGTGTGACTTTGAAACCGGTGAGCGCTGATGCTTTATTTGCTTAAATGA 315
Db      256 AAATTAATTTGAGCTCTGATGAACCCCTGTGGCTTATGATGCTTAAATGAATGAACAT 315
Qy      316 GCTACAAAAGATGAGCTTCAAGCAACCAAGTTCTTATGAGATGCTTGAAGGACAAAGACA 375
Db      316 GCTTGAAGGAGCTGGAACAAATGAAGAAAGTACTGACAGAAATTAATGCTTCAAGACA 375
```

Db	76	GCTGATGCAAGAACTCTTGGGAAAGCTATGAAAGGCTTGGGCAAGATGAGAGAGATC	135
Qy	136	ATATCAATCTTGGCTCAGAGAAAGTGTGAACAGAGAAAGTCAATCAAGCAATACAC	199
Db	136	CTGACTCTGTGTGACATCCGGAAGTATGCTCAGCCGACGAAATCTCTGCACTTTAAAG	195
Qy	196	GAAACCTAAGCCGACAACTTCTCAAGACTCTTGAACAAGAGCTCTCAAGATTTTGAG	255
Db	196	ACTCTGTGTGGCAGGGATCTTCTGATGACCTGAAATCAGAACTTACGAAAAATTGAA	255
Qy	256	AGAGCTATCTTGTGTGGACTCTTGAACCCGGGTGACGTGATGTCTTTATTTGGCTAATGAA	315
Db	256	AAATTTATTTGTGGCTCTGATGAAGAACCTCTCGCTTTATGTATGCTTATGAACCTGAACAT	315
Qy	316	GCTACAAAAGATGACTTACAGCAACCAAGTTCTTATGAAAGTTGCTTGCACAAGACA	375
Db	316	GCCTTGAAGGAGACTGGAACAAATGAAAAAGTACGACGAAATTAATGCTTCAAGACA	375
Qy	376	TCAACGAGCTGTCTTACGCGTAGGCAAGCTTACATGCTCGCTCAAGAAAGCTCTTGAA	435
Db	376	CTGAGAAACTGAGACCATCAACGAATTTATGAGAAAGATTTGGCTCAAGCTTGA	435
Qy	436	GAGACGTTGCTCACAACACTACCGGTGACTTCAGAAAAGCTTTTGTTCTCTTGTACC	495
Db	436	GATGACGTGGTGGGGACACTTCAGGGATCTCACAGCGAATGTTGGTGTCTCTCTCAG	495
Qy	496	TCATACAGGTCACAGAGAAATGATGAAACATGACATTGGCTTACGAAAGCTAACTG	555
Db	496	GCTPACAGACCCCTGATCTGTAATGTATGAAAGCTCAAGTTGAACAAGAGCTCAAGCT	555
Qy	556	GTCATGAGAAATCAAGACAAGCACTACATGATGAG--ATGTTATTGAATCTTG	612
Db	556	TTATTTCAAGCTGAGAACTTAATGGGGGACAGATGAAGAAAGTTATACCATCTTT	615
Qy	613	TCACAAGAAACAAAGCTCAGATCAATGCTACTTTAACCGTTACCAAGATGATAGGC	672
Db	616	GGAACACGAAGTGTCTCATTTGAGAAAGGTGTTGCAAGTACATGACTATACGGA	675
Qy	673	GAGAAATTTCTCAAGAGCTTTGAGAGAGAGATGATGACAAGTTCTTGCACTTTTG	732
Db	676	TTTCAATTGAGAAACATTTAAC--GCGAGCTTGCGCAATTTAGAGCAACTATC	733
Qy	733	AGGTCACCAATTCAGTCTTGACAAGACCAAGCTTATCTTGTGATGTTCTTGTTCA	792
Db	733	CTTGCTGTGTGAATCTTATGGAAGTATACCTGCTACCTTGCAAGAACCTCTATATAT	792
Qy	793	GCATCAACAAACCTGGAACCTGATAGAGAGACCTCAATATTTGATCAACAAGACT	852
Db	793	GCTATGAGGAGCTGGGACAGATATATACCTCTCAAGAGTCATAGTGTTCAGAGAGT	852
Qy	853	GAGATTGACTTGAAGCTCTTGTGAGAGAGTACACGCGCAGAAACAGACTCTTTGGAG	912
Db	853	GAGATTGACTGTTTAATCAACGAAAGGAGTTTGAAGAAATTTTGGCACCTGCTTTAT	912
Qy	913	AAAGCTATTTACAAAGACCTGCTGGAATTTACGAAAGATGCTGTGCACTTTGGGT	972
Db	913	TCCATGATTTAAGGAGATACATCTGGGACTATAAGAAAGCTTCTGTGCTCTCCGGA	972
Qy	973	GAAGATGA 980	
Db	973	GAAAGTGA 980	

```

RESULT 15
US-09-324-096A-5
; Sequence 5, Application US/09324096A
; Patent No. 632313
; GENERAL INFORMATION:
; APPLICANT: Tait, Jonathan
; APPLICANT: Brown, David
; TITLE OF INVENTION: ANNEXIN DERIVATIVE WITH ENDOGENOUS CHELATION SITES
; FILE REFERENCE: USFW-1-13841
; CURRENT APPLICATION NUMBER: US/09/324, 096A

```

;/ CURRENT FILING DATE: 1999-06-01  
;/ NUMBER OF SEQ ID NOS: 12  
;/ SOFTWARE: PatentIn version 3.0  
;/ SEQ ID NO 5  
;/ LENGTH: 981  
;/ TYPE: DNA  
;/ ORGANISM: Homo sapiens  
;/ FEATURE:  
;/ NAME/KEY: CDS  
;/ LOCATION: (1)..(981)  
US-09-324-096A-5

Query Match 8.3%; Score 101.6; DB 3; Length 981;  
Best Local Similarity 46.0%; Pred. No. 1.9e-19;  
Matches 418; Conservative 0; Mismatches 484; Indels 6; Gaps 2;

QY 76 GATGATGCTGAGCAATTGAGAACCGCTTTGAAAGATGGGGTACGAACGAGACTTGATC 135  
DB 76 GCTGATGCAAGAACTCTTCGGAAGGCTATGAAAGGCTGGGCAAGATGAGAGACATC 135  
QY 136 ATATCAATCTGGCTACAGAGAGTGTGAACAGAGAAAGTCATCAGGCAAGCATACAC 195  
DB 136 CTGACTCTGTTGACATCCGAGATTAATGCTCAGCGCCAGAAATCTCTGACGCTTTAAG 195  
QY 196 GAAACCTACGCGCAGAACCTTCTCAAGACTCTTGACAAGAGCTCTTAACGATTCGAG 255  
DB 196 ACTCTGTTGGCAGGGATCTTCTGATGACCTGAATCAGAACTAAGTGGAAATTTGAA 255  
QY 256 AGAGCTATCTTGTGTGAGACTTTGAACCCGTTGAGCGTGAATGCTTATTTGGCTAATGA 315  
DB 256 AATTAATGTGCTCTGATGAACCCCTCTCGGCTTATGATGCTTATGAATGAAACAT 315  
QY 316 GCTACAAAAGATGACTTGAACCAACAGTCTTATGGAAGTGTGCTGACAGAGACA 375  
DB 316 GCTTGAAGGAGCTGGAACAAATGAATTAATGCTGACAGAAATTAATGCTTGAAGACA 375  
QY 376 TCAACGACGCTGCTTCAAGCTAGGCAAGCTTACCATGCTCGTACAAGAAATCTTTGAA 435  
DB 376 CTTGAAGAACTGAGAGCATCAACAAAGTTTATGAAGAATATGGCTCAAGCTGAA 435  
QY 436 GAGAGCTTGTCTACCACTACCTACCTGACTTCAAGAACTTTGTTCTCTTTACC 495  
DB 436 GATGACGTGTGGGGGCACTTCAGGGTACTTACAGCGATGTTGGTGTCTCTTCAG 495  
QY 496 TCATACAGGTACGAAGGAGATGAATGATCATGACATTTGGCTAAGCAAGCTAAGCTG 555  
DB 496 GCTAACAGAGACCTGATGCTGGAATGTAGAACTCAAGTGAACAAGTGTCAAGGCT 555  
QY 556 GTCATAGAGAAAATCAAGACAGCACTACATGATGAGG--ATGTTATTAGAACTTG 612  
DB 556 TTATTTCAAGCTGAGAGACTTAATGGGGGACAGATGAAGAAAGTTTATCACCATCTT 615  
QY 613 TCACAGAAAGCAAGCTCAGATCAATGCTACTTTTAACTTAACTTAACTTAACTTAACT 672  
DB 616 GGAACAGAGTGTCTCTTGAAGAAAGTGTGACAAATGATACATGACTATATACGAG 675  
QY 673 GAGGAATTTCTCAAGGTCTTGAGAGAGATGATGACAAAGTTCCTTGCACTTTG 732  
DB 676 TTTCAATTTGAGAAACCATTTGACC--GCGAAGCTTCTGGCAATTTAAGCAACTATC 732  
QY 733 AGGTCAACCATTCAGTGTCTTGACAGACAGAGCTTACTTGTGATGCTTCTGCTCA 792  
DB 733 CTGCTGTGTGTGAATCTATGGAATATCTGCTTACTTGCAGAGACCTCTATATAT 792  
QY 793 GCAATCAACAAAATGGAAGTGAATGAGAGCACTCACTAATTTGTGACACAGAGCT 852  
DB 793 GCTATGAGGAGAGCTGGGACAGATGATCAATACCTCATCAGAGTCAATGTTTCAGAGAT 852  
QY 853 GAGATTGACTTGAAGTCAATTTGAGAGAGTACAGAGCAAGCAATTCCTTTGAG 912  
DB 853 GAGATTGATCTGTTTAACTCAAGAGAGAGTTTGAAGAAATTTTGCACCTCTCTTAT 912  
QY 913 AAGCTATTACAAAGCACTCGTGAAGATTACGAGAAGATGCTCTGCACTTCTGGAT 972

DB 913 TCCATGATTAAGGAGATATCATCTGGGGACTATAGAAAGCTCTTCTGCTGCTCCGGA 972  
QY 973 GAAAGTGA 980  
DB 973 GAAAGTGA 980

Search completed: August 22, 2005, 03:09:22  
Job time : 260 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2005, 02:35:40 ; Search time 2737 Seconds

(without alignments)  
2919.697 Million cell updates/sec

Title: US-10-690-564-1

Perfect score: 1230

Sequence: 1 ccacgcgtccgaactaa.....aaaaaaaaaaaaaaaa 1230

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/prodata/2/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/prodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/prodata/2/pubpna/US09A\_PUBCOMB.seq:\*

10: /cgn2\_6/prodata/2/pubpna/US09B\_PUBCOMB.seq:\*

11: /cgn2\_6/prodata/2/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq:\*

13: /cgn2\_6/prodata/2/pubpna/US10A\_PUBCOMB.seq:\*

14: /cgn2\_6/prodata/2/pubpna/US10B\_PUBCOMB.seq:\*

15: /cgn2\_6/prodata/2/pubpna/US10C\_PUBCOMB.seq:\*

16: /cgn2\_6/prodata/2/pubpna/US10D\_PUBCOMB.seq:\*

17: /cgn2\_6/prodata/2/pubpna/US10E\_PUBCOMB.seq:\*

18: /cgn2\_6/prodata/2/pubpna/US10F\_PUBCOMB.seq:\*

19: /cgn2\_6/prodata/2/pubpna/US10G\_PUBCOMB.seq:\*

20: /cgn2\_6/prodata/2/pubpna/US10H\_PUBCOMB.seq:\*

21: /cgn2\_6/prodata/2/pubpna/US10I\_PUBCOMB.seq:\*

22: /cgn2\_6/prodata/2/pubpna/US10J\_PUBCOMB.seq:\*

23: /cgn2\_6/prodata/2/pubpna/US11A\_PUBCOMB.seq:\*

24: /cgn2\_6/prodata/2/pubpna/US11\_NEW\_PUB.seq:\*

25: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq:\*

26: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1230	100.0	21	US-10-690-564-1	Sequence 1, Appl 1
2	635	51.6	635	US-09-770-149-655	Sequence 655, App
3	497.8	40.5	1184	US-10-767-795-4242	Sequence 4242, Ap
4	479.6	39.0	1145	US-10-425-114-7439	Sequence 7439, Ap
5	479.6	39.0	1145	US-10-425-114-29928	Sequence 29928, A
6	479.6	39.0	1393	US-10-424-599-1190	Sequence 1190, Ap
7	473.4	38.5	1058	US-10-219-220-218	Sequence 218, App

	8	471.8	38.4	2588	17	US-10-393-840-28	Sequence 28, Appl
C	9	470	38.2	487	10	US-09-770-961-765	Sequence 765, App
	10	468.6	38.1	1293	14	US-10-219-220-17	Sequence 17, Appl
	11	468.6	38.1	1293	17	US-10-393-840-109	Sequence 109, App
	12	450.4	36.6	1145	18	US-10-425-114-10772	Sequence 10772, A
	13	430.8	35.0	1331	18	US-10-424-599-70457	Sequence 70457, A
	14	364.2	29.6	789	14	US-10-219-220-20	Sequence 20, Appl
	15	364.2	29.6	789	17	US-10-393-840-26	Sequence 26, Appl
	16	364.2	29.6	789	17	US-10-393-840-112	Sequence 112, Appl
	17	358.6	29.2	908	18	US-10-425-114-10228	Sequence 10228, A
	18	353	28.7	1235	18	US-10-425-114-35285	Sequence 35285, A
	19	346.6	28.2	1235	18	US-10-425-114-23965	Sequence 23965, A
	20	346.6	28.2	1243	18	US-10-425-114-23365	Sequence 23365, A
	21	346.6	28.2	1285	18	US-10-425-114-26803	Sequence 26803, A
	22	346.6	28.2	1385	19	US-10-767-701-13525	Sequence 13525, A
	23	346.6	28.2	1972	20	US-10-425-115-2178	Sequence 2178, Ap
	24	338.6	27.5	1337	19	US-10-767-701-14545	Sequence 14545, A
	25	338.6	27.5	1496	20	US-10-425-115-63693	Sequence 63693, A
	26	337	27.4	1300	20	US-10-739-930-2276	Sequence 2276, Ap
C	27	336.2	27.3	1476	19	US-10-437-963-58743	Sequence 58743, A
	28	327.2	26.6	602	19	US-10-021-323-12115	Sequence 12115, A
	29	327	26.6	591	19	US-10-021-323-15046	Sequence 15046, A
	30	326	26.5	589	19	US-10-021-323-13465	Sequence 13465, A
	31	324.8	26.4	1292	19	US-10-437-963-78307	Sequence 78307, A
	32	322.2	26.2	600	19	US-10-021-323-15621	Sequence 15621, A
	33	315.8	25.7	586	19	US-10-021-323-15526	Sequence 15526, A
	34	311.2	25.3	1437	18	US-10-424-599-30062	Sequence 30062, A
	35	296.4	24.1	1497	18	US-10-424-599-88585	Sequence 88585, A
	36	284.2	23.1	1153	18	US-10-425-114-7980	Sequence 7980, Ap
C	37	284.2	23.1	1194	20	US-10-739-930-3735	Sequence 3735, Ap
	38	282.6	23.0	570	19	US-10-021-323-13010	Sequence 13010, A
	39	277.8	22.6	1172	18	US-10-425-114-13073	Sequence 13073, A
	40	277.8	22.6	2336	18	US-10-424-599-137007	Sequence 137007, A
	41	277	22.5	1113	18	US-10-424-599-130903	Sequence 130903, A
	42	276.2	22.5	2612	18	US-10-424-599-68473	Sequence 68473, A
	43	273.2	22.2	1320	18	US-10-424-599-137005	Sequence 137005, A
	44	271	22.0	608	19	US-10-021-323-8405	Sequence 8405, Ap
	45	270.6	22.0	2132	17	US-10-393-840-27	Sequence 27, Appl

#### ALIGNMENTS

RESULT 1

US-10-690-564-1

Sequence 1, Application US/10690564

Publication No. US20050089872A1

GENERAL INFORMATION:

APPLICANT: KOREA KIMHO PETROCHEMICAL CO., LTD

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ANNEXINS FROM PLANTS

FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/10/690,564

CURRENT FILING DATE: 2003-10-23

NUMBER OF SEQ ID NOS: 4

SOFTWARE: Kopatentlin 1.71

SEQ ID NO 1

LENGTH: 1230

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-10-690-564-1

Query Match	Score	DB	Length
Best Local Similarity	100.0%	1230;	
Matches 1230;	Conservative	0;	Mismatches 0;
		Indels	0;
		Gaps	0;

QY 1 CCACGCGTCGGAACCTAAAGTAGAGAAAATGCGACCTTAAGTTCTGATTC 60

DB 1 CCACGCGTCGGAACCTAAAGTAGAGAAAATGCGACCTTAAGTTCTGATTC 60

QY GTTCTGCTCTCTTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGGATAC 120

DB GTTCTGCTCTCTTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGGATAC 120

OY	121	AACGAGACTTGATCATATCAATCTTGGGTCCAGAAGTGCCTGAACAAGAAAGTCATC	180
Dd	121	AACGAGACTTGATCATATCAATCTTGGGTCCAAGAAGTGCCTGAACAAGAAAGTCATC	180
OY	181	AGCGAAGCATCACGAAACCTAACGCGCAAAGACTCTTCACAGACTCTTGAACAAGAGCTC	240
Dd	181	AGCGAAGCATCACGAAACCTAACGCGCAAAGACTCTTCACAGACTCTTGAACAAGAGCTC	240
OY	241	TCTAA CGATTTCGAGAGACTATCTTTGTGTGAACTCTTGAACCCGGTAGGCCGTATGCT	300
Dd	241	TCTAA CGATTTCGAGAGACTATCTTTGTGTGAACTCTTGAACCCGGTAGGCCGTATGCT	300
OY	301	TTATTGGCTAATGAAGCTACAAAAAGATGSACTTCAAGCAACAAAGTTCTTATGSAAGTT	360
Dd	301	TTATTGGCTAATGAAGCTACAAAAAGATGSACTTCAAGCAACAAAGTTCTTATGSAAGTT	360
OY	361	GCTTGACAAGGACATCAACGAGCTGTCAACGTAAGCAAGCTTACATCACTGCGTAC	420
Dd	361	GCTTGACAAGGACATCAACGAGCTGTCAACGTAAGCAAGCTTACATCACTGCGTAC	420
OY	421	AAGAAGTCTCTTGAAGAGACCCTTCTACCACTACCGGTGACTTGAGAAAGCTTTTG	480
Dd	421	AAGAAGTCTCTTGAAGAGACCCTTCTACCACTACCGGTGACTTGAGAAAGCTTTTG	480
OY	481	GTTTTCTCTTGTATCTCATACAGATACGAAGAGATGAAGTACATGACATTTGGCTAAG	540
Dd	481	GTTTTCTCTTGTATCTCATACAGATACGAAGAGATGAAGTACATGACATTTGGCTAAG	540
OY	541	CAGAAGGCTAACCTGGTCCATGAGAAATCAAGSACAAAGCTACATAGATGAGATGTT	600
Dd	541	CAGAAGGCTAACCTGGTCCATGAGAAATCAAGSACAAAGCTACATAGATGAGATGTT	600
OY	601	ATTAGATCTTGTCCACAAGAAAGCAAGCTCAGATCAATGCTTCTTAAACGTTAACCA	660
Dd	601	ATTAGATCTTGTCCACAAGAAAGCAAGCTCAGATCAATGCTTCTTAAACGTTAACCA	660
OY	661	GATGATCATGCGGAGAGAAATTTCTAAGAGCTTTGAGAAAGAGATGATGACAAAGTTG	720
Dd	661	GATGATCATGCGGAGAGAAATTTCTAAGAGCTTTGAGAAAGAGATGATGACAAAGTTG	720
OY	721	CTTGCACCTTTGAGGTCAAACCATGAGCTTGAACAACGACACTTACCTTGTGCAT	780
Dd	721	CTTGCACCTTTGAGGTCAAACCATGAGCTTGAACAACGACACTTACCTTGTGCAT	780
OY	781	GTTCTTCGTTCAGCATCAACAACCTGGACCTGATGAAGAGACCTCACTAGATTGTG	840
Dd	781	GTTCTTCGTTCAGCATCAACAACCTGGACCTGATGAAGAGACCTCACTAGATTGTG	840
OY	841	ACCAACAAGACTGAGATTGACTTGAAGGTCATTTGAGAGAGATACCAACGCGAGAAACAGC	900
Dd	841	ACCAACAAGACTGAGATTGACTTGAAGGTCATTTGAGAGAGATACCAACGCGAGAAACAGC	900
OY	901	ATTCTTTGGAGAAAGCTATTAACAAGACATCGTGAGATTAGAGAAAGTGTGCTTC	960
Dd	901	ATTCTTTGGAGAAAGCTATTAACAAGACATCGTGAGATTAGAGAAAGTGTGCTTC	960
OY	961	GCACTTTCGAGAGATGATGCTTATCATCAATCTCCACAGAGAAACATAAGCTGC	1020
Dd	961	GCACTTTCGAGAGATGATGCTTATCATCAATCTCCACAGAGAAACATAAGCTGC	1020
OY	1021	TCTACAGCTTCTGTTATCTCTTATCTCCCTCTCTCTCTCTTGTGATGAGTTTCAATCGTT	1080
Dd	1021	TCTACAGCTTCTGTTATCTCTTATCTCCCTCTCTCTCTCTTGTGATGAGTTTCAATCGTT	1080
OY	1081	TGATTTTGTTTTACAAAACCTTGTGTTTCTGTGTGTGTTTGAAGTTCTTAAATPA	1140
Dd	1081	TGATTTTGTTTTACAAAACCTTGTGTTTCTGTGTGTGTTTGAAGTTCTTAAATPA	1140
OY	1141	TGCAAAAGAGAGACAGAGAACCAAGTGTGCTCTTAAGTTATATATATATGAAGAG	1200
Dd	1141	TGCAAAAGAGAGACAGAGAACCAAGTGTGCTCTTAAGTTATATATATATGAAGAG	1200
OY	1201	CATTGGCCTAAAAAAAAAAAAAAAAAAAAA 1230	

[illegible]

Db	361	TCACGCGAGCTGTTCAACGTAGGCAAGCTTACATGCTCGGTACAAABAATCTCTTGAA	420
Qy	436	GAGAGCTTGTCTCAACAACACTACCGGTGACTTCAGAAAAGCTTTGTGTTCTCTTGTTACC	495
Db	421	GAGAGCGTTGTCTTACACCACTACCGGTGACTTCAGAAAAGCTTTGTGTTCTCTTGTTACC	480
Qy	496	TCATACAGGTACGACGAGATGATGAACATGACATTGGCTTAAGCAAGAGCTTAAGCTG	555
Db	481	TCATACAGGTACGACGAGATGATGAACATGACATTGGCTTAAGCAAGAGAGCTTAAGCTG	540
Qy	556	GTCCTATGAGAAAATCAAGGACAAGCACTACATGATGATGAGATGTTATTAGATCTTGTCC	615
Db	541	GTCCTATGAGAAAATCAAGGACAAGCACTACATGATGATGAGATGTTATTAGATCTTGTCC	600
Qy	616	ACAAGAAGCAAAAGCTAGATCAATGCTACTTTTAA	650
Db	601	ACAAGAAGCAAAAGCTAGATCAATGCTACTTTTAA	635

```

RESULT 3
US-10-767-795-4242
: Sequence 4242, Application US/10767795
: Publication No. US2004018130A1
: GENERAL INFORMATION:
: APPLICANT: Kovalic, David K.
: APPLICANT: Cao, Yongwei
: APPLICANT: Zhou, Yihua
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
: FILE REFERENCE: 38-21(53534)B
: CURRENT FILING DATE: 2004-01-30
: NUMBER OF SEQ ID NOS: 117596
: SEQ ID NO 4242
: LENGTH: 1184
: TYPE: DNA
: ORGANISM: Gossypium hirsutum
: FEATURE:
: OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C101_1
: US-10-767-795-4242

```

Query Match	40.5%	Score 497.8	DB 19	Length 1184
Best Local Similarity	70.7%	Pred. No. 2.3e-110		
Matches 691	Conservative 0	Mismatches 282	Indels 4	Gaps 2
Qy	12	AAACCTAAAAAGTAGAAGA-AAATGGGCACTTTAAGGTTTCGATTCTGTCCTGCTC	70	
Dp	47	AAEAAATAAAGGAAGAAAGCAATGGCCACTTTACAGTCCCAAGCAAGTTCTTGGG	106	
Qy	71	CTTCTGATGATGCTGAGCAATTGAGAACCCGCTTTTGAAGATGGGGTTCGAACGAGACT	130	
Dp	107	TGTTTGAGAGATTGGAAACGCTTAAGAAAGCTTTTCAGATGGGGAATATATGAGGCT	166	
Qy	131	TGATCATATCATCTTTGGCTCAGACGAAGTGTGAAACAGAGAAAGTCATCAGGCAAGAT	190	
Dp	167	TATCATATGATATATTGGGTACAGAAATGCCGACACAGAACTTGATTGGAAAAACT	226	
Qy	191	ACCACGAACCTTACGGCGAAGACCTTCTCAAGACTTTGACAGAGAGCTCTTAACGATT	250	
Dp	227	ACGCTGAACCTTATGAGAGAGATCTCTCAAGGACATAGACAGAGAGCTTCGATGACT	286	
Qy	251	TTCGAGAGAGCATTTGTGTGACCTTTGAAACCCGGAGACGGATGCTTATTGGCTA	310	
Dp	287	TTGAGAGGCTGGTTTGCTTTGGGCTCTTGATCTGCTGAACGGATGACCTTTTGGCTA	346	
Qy	311	ATGAGCTACAAAAAGATGACATTTCAAGCAACAAAGTCTTATGAAAGTTGCTTGACAA	370	
Dp	347	ATGAAAGCCAAAGGTGACCTTCAAGCAATCAGTCTTATGAAATAGCTTGACAA	406	
Qy	371	GGACATCAACGCGAGCTGTTACGCTTAGGCAAGCTTACATGCTGCTACAAAGATCTC	430	
Dp	407	GGTTTGCAACCAACTCTTCAAGCAAGGACGGCTTATCATAGCTCGTTATTAAGAAAGTCG	466	

Oy	431	TTGAAGAGAGCTTGTCTCCACCACTACCGGTGACTTCAGAAAGCTTTGGTTTCTCTTG	490
Db	467	TTGAAGAGAGATGGTGTCTCATCCACGACTGGCCGCTTCGTAAGCTCTCTCACTCTTAG	526
Oy	491	TTACCTCATACAGGTACGAGAGAGATGAGTGAACATGACATTTGGCTTAAGCAAGCTA	550
Db	527	TGAATTATACAGTATATGAGGAGAGAGAGTGAACATGATCTGGCCGAAAAACAGGCGA	586
Oy	551	AGCTGGTCCATGAGAAAATCAAGGACCAAGCACTACAAATGATGAGATGTATTAAGAACT	610
Db	587	AGTTGCTTCATGAGAAAATTTACAGCAAAAGCTTACAGTATGACGATGTCAATAGGGTTT	646
Oy	611	TGTCCACAAGAGGAAAGCTCAGATCAATGCTACTTTAAACCTTACCAGATGATCATG	670
Db	647	TGGCTACAAAGAGCAAGGACAGATCATGATGAACTTCGAAATCACTAACAAAATTAATATG	706
Oy	671	GCGAGAAATTTCTCAAGATCTTGAGGAAGAGATGATGATGACAAAGTCTCTTGACATTT	730
Db	707	GAAATGCAATTAACAAAGACTT---GAAAGCTGATCTTAAGATGAGTTCTTTCACATAC	763
Oy	731	TGAGGTCAACCATTTCAAGTCTTGAACAAGCCAGACTTTTACTTTGTGATGTTCTTGCTT	790
Db	764	TAAAGTCCACAGTGAAGTGGTGGTCTATCCGAGAAAAGTATTTTGAAGAAGTTCTTGCC	823
Oy	791	CAGCAATCAACAAAACCTGGAATGTATGAAGAGCACTCACTAGAAATTTGTACCAACAAG	850
Db	824	TAGCAATCAATAGACGGAACCGATTAAGAGCTCTTACTTAGAGTTGTTGCACTAGGG	883
Oy	851	CTGAGATTGACTTGAAGGTCAATTTGGAAGAGTACACGCCAGGAACAGCAATTCCTTTGG	910
Db	884	CTGAGTTGATCTTAAGATCATAGCAGATGATGTCACGACGAAGAACAGTGTCCCACTGA	943
Oy	911	AGAAAGCTATTACAAAGACACTTGTGGAATTAACGAAGATGTCTGTGCACTTCTCG	970
Db	944	CTGTGCACTATGTCAAGACACTTCATGGAAGCTATGAATAATGCTCTGTGTACTTGCAAG	1003
Oy	971	GTGAAGTATGATGCTTAA 987	
Db	1004	GACATGTGAGAAATGGA 1020	

```

RESULT 4
US-10-425-114-7439
; Sequence 7439, Application US/10425114
; Publication NO. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7439
; LENGTH: 1145
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700661911_FLI
US-10-425-114-7439

```

Query Match	39.0%	Score 479.6	DB 18	Length 1145
Best Local Similarity	69.0%	Pred. No. 3.3e-125		
Matches 672	Conservative	0	Mismatches 259	Indels 3
			Gaps	1
14	ACACTAAGTAGAAGAAAAATGCGACCTCTTAAGCTTTCTGATTCCTGCTCTT	73		
34	AAACAAAAACGAGACACATGCGACACTTAAGCTTCTGACCTCTTCCCCCGGTG	93		



Oy	74	CTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAAGATGGGGATACGAGAGACTTGA	133
Db	94	CAGATGATGTGAGAGCTCCGAAAACCTTCTCAGTTGGGGAACTTAAACGAGAGCTGA	153
Oy	134	TCATATCAATCTTGGGCTCCAGAAAGTCTGAACAGAGAAAGTATCAGGCAGACATACC	193
Db	154	TTGATATGATCTTGGCTTCATAGGAATGCTGCTCAGAGAGACTTATCAGAGAACTTATG	213
Oy	194	ACGAAACTTACGGCGAAGACCTTCTCAGACTCTTGACAGAGAGCTCTTAAACGATTTTCG	253
Db	214	CCGAGACTTATGAGAGAAAGATCTCTCAAGGCTTGAGACAAAGAACTCACGAGTATTTTG	273
Oy	254	AGAGAGCTATCTTGTGTGTGACTCTTGAACCCGGTGACGTGATGCTTTATTTGGCTATG	313
Db	274	AGAGGCTGTTCATCTTTTGGACACTTGATTCGTCTGAACGTGATGCAATTTTGGCAACG	333
Oy	314	AAGCTACAAAAGATGAGACTTCAAGCAACCAAGTCTTATGGAAGTGTCTTGACACAAGA	373
Db	334	AGCGACTTAAAAATGAGCTTCAAGCAATCAGGTTCTGGTGGAAATAGCTGCACTAGGT	393
Oy	374	CATCAACGACGCTGCTTCAACGCTAGGCAAGCTTACCATGCTCGCTACAGAAAGTCTTTG	433
Db	394	CCTCTGAACAATGTTTGGCTGGAGGAAGCTTACCATGTTCTTATATGAAGTCTCTGG	453
Oy	434	AAGAGACGTTTGCTCACACACTACCGGTGACTTCAAGAAAGCTTTTGTTCTTGTTA	493
Db	454	AGGAGATGTTGCTCATCACACAAACGAGAGCTTCCGTAACCTCATATCACTCTCTGTTTA	513
Oy	494	CCTCATCAGATGCAAGAGGATGAAAGTAAACATGACATGAGTGGCTTAAGCAAGACTTAAGC	553
Db	514	GTTCTTATGCAATTTGAAGGATAGAGTCAACTTGAACCTGGCAAAAAGTGAAGCAAAAT	573
Oy	554	TGCTCCATGAGAAAATCAAGGACAAAGCACTCAATGATGAGATGTTATTAAGAACTCTGT	613
Db	574	TGCTGATGAGAGATTTCAACAAAGGCTTATATGATGACGACTTACAGAGATTTTGG	633
Oy	614	CCACAAAGACAAAGCTCAGATCAATGCTACTTTTAAACGTTTACCAGATGATATATGGCG	673
Db	634	CCACAAAGACCAAGACAGATTAAGTCACTTTGATATCATCAAAAGATGCAATTTGGAC	693
Oy	674	AGGAAATTCACAGAGCTTGAGGAAGGATGATGATGACAAAGTTCCTTGCACTTTTGA	733
Db	694	AAGTATCAACAAAGGACT---GAGGCTGATCCAAAGAGAGATTTCTGTCTTATCTAA	750
Oy	734	GGTCAACACTTACAGTCTTGACAAAGCCAGACTTACTTTGTGATGATTTCTTGTCAG	793
Db	751	GAGCAACTGTAAGTGTGTTGATCCGTCGGAAGAGTACTTGAGAAAGTGTGTGCTTTGG	810
Oy	794	CAATCAACAAACTCGAATCTGATGAAGAGGCACTCACTAAGATTTGTGAACCAAGAGCTG	853
Db	811	CGATCAACAAAGCGAGGAACGATGAAGGAGCCCTTACAAAGTGTGTGCCCAAGGGCTG	870
Oy	854	AGATTGACTGAAGGTCAATTGAGAGAGATACCAAGGCGAGAAACGACTTCTTTGGAGA	913
Db	871	AGGTTGATTTGAAAAACATAGCAATGATGATACAAAGAGGAGCAAGTGTCCCTCTTGAAC	930
Oy	914	AAGCTATTAACAAAGACACTCGTGAGATTACGAGAAAGTGTGCTGCACTTCTCGGTG	973
Db	931	GTCCTATTTGTTAAGGACACCAACCGGTGACTATGAGAAATGTTGTGCGCTTTTAAGAC	990
Oy	974	AAGATGATGCTTAA 987	
Db	991	ATGATGATGCTTGA 1004	

RESULT 5  
US-10-425-114-29928  
; Sequence 29928, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingtong  
; APPLICANT: Zhou, Yihua

```

1  APPLICANT: Kovacic, David K.
2  APPLICANT: Screen, Steven E
3  APPLICANT: Tabaska, Jack E
4  APPLICANT: Cao, Yongwei
5  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
6  TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
7  FILE REFERENCE: 38-21(53133)B
8  CURRENT APPLICATION NUMBER: US/10/425,114
9  CURRENT FILING DATE: 2003-04-28
10 NUMBER OF SEQ ID NOS: 73128
11 SEQ ID NO 29928
12 LENGTH: 1145
13 TYPE: DNA
14 ORGANISM: Glycine max
15 FEATURE:
16 OTHER INFORMATION: Clone ID: UC-GMROPTC033E01_FLI
17 US-10-425-114-29928

```

Query Match	39.0%;	Score 479.6;	DB 18;	Length 1145;
Best Local Similarity	69.0%;	Pred. No. 3.3e-125;		
Matches 672;	Conservative 0;	Mismatches 299;	Indels 3;	Gaps 1

OY	14	ACACTAAAGATGAAAGAAATATGCGACTCTTAAGTTCTTGATTCGTGTCCTGCTT	73
Db	34	AAACAAAAACAGAACCAATATGGGACATTAAGTTCTCAGCTCTTCCCGCTTG	93
OY	74	CTGATGATGCTGAGCAATTTGAGAACCGCTTTTGAAGATATGGGGTACGAACGAGCTTGA	133
Db	94	CAGATGATTTGACGACGCTCCGAAAAGCTTCTCAGGTTGGGGAACTAACGAGAGCTGGA	153
OY	134	TCATATCATCTTGGCTCACAGAAAGTGTGAACAGAGGAAGTCAACAGCAAGCATAC	193
Db	154	TTGATATCATCTTGGCTCATAGGAATCTGCTCAGAGGAAGCTTATCAGAGAAACTTATG	213
OY	194	AAGAAACCTAGCGGGAAGACCTCTCAAGACTCTTGAACAAGAGCTCTCAACATTTG	253
Db	214	CCGAGACTTATGAGAGAAATATCTCTCAAGGCTTGGACAAAGAACTCACAGATGATTTG	273
OY	254	AGAGAGCTATCTTGTGTGAGACTCTTGAACCCGAGCGATGATCTTTATTTGAGTAAATG	313
Db	274	AGAGGCTGTTTCATCTTTTGGACACTGTAATTCGTGGAACGTGATTCATTTTGGCAACG	333
OY	314	AAGCTACAAAAGATGAGACTTCAAGCAACCAATTTCTTATGGAAGTTGCTTGCACAGGA	373
Db	334	AGCGACATTAATAATGACTTCAAGCAATCAGTCTTGATGAAATACCTGCACTAGGT	393
OY	374	CATCAAGCAGACTGTTTCAACGCTAGGCAAGCTTACATGCTCGTACAAAGATCTCTTG	433
Db	394	CCTCTGAACAAATGTTTGTCTCGAGGAAGGCTTTACATGTTCTTTATAGAAGTCTCTGG	453
OY	434	AAGAGGAGTGTGCTCAACCACTACCGGTGACTTCAGAAAGCTTTTGGTTTCTTTGTTA	493
Db	454	AGGAGGATTTGCTCATACACACAGAGACTTCCGTAAGCTACTACTCTCTGGTTA	513
OY	494	CCTCATACAGTACGAAAGAGATGAATGAACATGACATTTGGCTTAAGCAAGAGCTAAGC	553
Db	514	GTTCTTATCATATGAAGAGATGAGTCAACTTGAACCTGGCAAAAACCTGAGGCATAAT	573
OY	554	TGCTTCATGAGAAATCAAGGACAAAGCACTACATGATGAGAGATGTTATTAAGATCTTGT	613
Db	574	TGCTGCAATGAAMAGTTTCAACAAAGGCTTTTAAGATGACGACTTCATCAGAGATTTTGG	633
OY	614	CCACAAGAGCAAGAGCTCAGATCAATGCTACTTTTAACCGTTTACCAAGATGATCATGGCG	673
Db	634	CCACAGAGGACAGACACAGATTATGCTTACTTTGAATCACTACAAAGATGATTTGGAC	693
OY	674	AGGAAATTTCTCAAGGCTTTGAGAGGAGATGATGATGACAAAGTTCTTGCACATTTTGA	733
Db	694	AAAGATATCAACAAAGGAACT---GAAGGCTGATCCAAAGAGCGAGTTCTGCTTTATCTAA	750
OY	734	GATCAACCATTCAGTGTGTTGACAAAGACCAAGAGCTTTACTTTGTGCATGTTCTTGGTCAAG	793
Db	751	GAGCAACTGTTAAAGGCTTGGATCCGCTCCGAGAGATGATCTCGAAGAAAGTGTGTTGCTTGG	810

Qy	79	CAATCAACAACCTGGAACTGATGAGAGACACTAGATATTGTGACCAAGAAGCTG	853
Db	811	CGATCAACAACGAGGAACGATGAAGAGACCTTTACAAGAGTGTGGCCACAAGGGCTG	870
Qy	854	AGATTGACTTGAAGGTCAATTGGAGAGAGTACACAGCGCAGAAACAGACTTCTTGGAGA	913
Db	871	AGCTTGATTTGAAAAACATAGACGATGATATCAAGAAGAGAGCGATGCTCCCTTTGAAAC	930
Qy	914	AAGCTATTACAAAGACACTCGTGAGATTACGAGAAAGTCTCGTGCACCTTCTCGGTG	973
Db	931	GTGCTATTGTTAAGGACACCAACCGGTGACTATGAGAAATGCTGTGGCCTTTTAGGAC	990
Qy	974	AAGTGAATGCTTAA	987
Db	991	ATGATGATGCTTGA	1004

```

RESULT 6
US-10-424-599-1190
: Sequence 1190, Application US/10424599
: Publication No. US20040031072A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa Thomas J
: APPLICANT: Kovalic David K
: APPLICANT: Zhou yihua
: APPLICANT: Cao Yongwei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53223)B
: CURRENT APPLICATION NUMBER: US/10/424,599
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 285684
: SEQ ID NO 1190
: LENGTH: 1393
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT3847_101073C.1
US-10-424-599-1190

```

Query Match	Similarity	Score	DB	Length
Query Match	69.0%	479.6	DB 18	1393
Best Local	69.0%	Pred. No. 3.7e-125		
Matches	672	Conservative	0	Mismatches 299; Indels 3; Gaps 1;
QY	14	ACACTAAAGTAGAAGAAAAATGGCACTTAAAGTTCTGATTTCTGTCGTCCTT	73	
Db	113	AAACAAAACAGAACAAACATGGGCACTTAAAGTTCTCAGCTTTTCCCCCGTTG	172	
QY	74	CTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGCGGTTCGAAACGACGACTTGA	133	
Db	173	CAGATGATTTGGAGAGCTCCGAAAGCCCTTCTCAGTTGGGGAACTAACACGAGGCTGGA	232	
QY	134	TCATATCAATCTTGGCTCACAAGAAAGTGTCTGAACAGAGAAAGTACATCAGGACAGCATAC	193	
Db	233	TTGATGATCTTGGCTCATAGAAAGTGTCTCACAAGAAAGCTTATCAGAAACTTATG	292	
QY	194	ACGAAACCTACGCGCAGACCTTCTCAAGACTTTGACAGAGAGCTCTTACAGATTTG	253	
Db	293	CCCAGACCTTAGGAGAAATCTCCTCAAGGCTTGGACAAAGAACTCAGAGTGAATTTG	352	
QY	254	AGAGAAGCATCTTGTGTGGACTCTGGAACCCGCGTAACGATGACTTATTTGGCTATG	313	
Db	353	AGAGCGCTGTCATCTTTTGGACACTTGATTTGCTGGAACGTAAGCAATTTTGGCAACG	412	
QY	314	AAGCTACAAAAGATGACTTCAAGCAACCAAGTCTTATGGAAGTTGCTTGCACAAAGA	373	
Db	413	AGGGAGCTAAAAAATGCACTTCAAGCATCAAGTTCTGTGTGAAATATGCTGCACATAGT	472	
QY	374	CATCAACGCACTGCTTACACGTAGGCAAGCTTATCATGCTCGCTACAAAGTCTCTTG	433	
Db	473	CCTTGAAACAATTTGTTCTCTCGAGGAAGGCTTACCAATGTTCTTATTAAGAAAGTCTCGG	532	

OY	434	AAGGAGCGTGTGCACCAACACTACCGGTGACTTCAGAAAGCTTTGGTTTCTCTGTGA	493
Db	533	AGGAGGATGTTGTCTCATCAACAACAAGACTTCCGTAAAGCTATCTACTCTCTGTGA	592
OY	494	CCTCATACAGGTATCGAAGAGATGAATGAACATGACATTGGCTTAACAAGAACTAAAC	553
Db	533	GTTCTTATCGATATGAAAGAGATAGGTCAACTTGACCTCGCAAAAACTAGAGCAAAAT	652
OY	554	TGCTTCATGAGAAAATCAAGACAAGCACTACAAATGATGAGATGTTATTGAATCTTGT	613
Db	653	TGCTGATGAGAAATTTCAACAAGGCTTATATGATGACGACTTATCAGATGTTTGG	712
OY	614	CCAAGAAGCAAGAGCTCAGATCAATGCTACTTTTAACGTTACCAAGATGATATGCGC	673
Db	713	CCACAAGAGAGAGACAAGATTATCTCTACTTGAATCACTAAAGAAATGCAATTTGGAC	772
OY	674	AGGAAATTCCTCAAGACTTGTGAGGAGAGAGATGATGATGACAAGTTCCTTGCACTTTGA	733
Db	773	AAGATATCAACAAGACT---GAAGCTGATCCAAAGACAGATTCTGTCTTTACTTA	829
OY	734	GGTCAACATTCACTGTGCTTGAACAAGACAAGCTTTACTTTTGATGATGTTCTTCGTTCA	793
Db	830	GAGCAACTGTGAATGTGCTTGATCCGTCCGAGAAAGTACTTCGAAAGSTTGTGCTTGG	889
OY	794	CAATCAACAAAACCTGGAATGATTAAGAAGACCTCACTAATTTGTATCAACAAGACTG	853
Db	890	CGATTCACAAAGCGAGGAACCGATATMAAGAGCCCTTAAAGAAGTGSTGCCACAAGGGCTG	949
OY	854	AGATTGACTTGAAAGSTCAATTGAGAGAGATTAACGCGCAGAAACAGCAATTCCTTTGAGA	913
Db	950	AGCTTGATTTGAATAAAACCTAGCAATGATGATCAAAAGAGAGACAGTGTCCCTCTTGAAC	1009
OY	914	AAGCTATTACCAAAAGACACTGTGGAATTAAGGAAGATGCTGTGCACTTCTCGGTG	973
Db	1010	GTCCTATTGTTAAAGACACCAACCGGTGACTATGAAAAATGCTGTGGCGCTTTTGAAGAC	1069
OY	974	AAGATGATGCTTAA	987
Db	1070	ATGATGATGCTTGA	1083

```

RESULT 7
US-10-219-220-218
; Sequence 218, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Laabam, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of plant development
; FILE REFERENCE: 11000.1022c1
; CURRENT APPLICATION NUMBER: US/10/219.220
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 218
; LENGTH: 1058
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-219-220-218

```

Query Match	38.5%	Score 473.4	DB 14	Length 1058
Best Local Similarity	69.0%	Pred. No. 1.8e-123		
Matches	664	Conservative	0	Mismatches 296; Indels 3; Gaps 1
Oy	25	AGAGAAAAATGGGAGACTTAAAGTTTCATCTGTCTCGCNCCTTGATGATGATCT	84	
Db	63	AGAGAGAAATGGGAGACTATCGCGGTGCACCGTCGGTTCCTCTCGCGCTGAGAGATCC	122	
Oy	85	GAGCAATTGACACCGCTTTGAAGATGGGATACGAACGAGACTTGATCATATCAATC	144	

```
Db 123 GAGCAGCTCCAAAAAGCTTTCGAGAGATGGGAAACGAATGAATGATCATATCCATA 182
QY 145 TTGGCTCAGAGAGTGTGAACAGAGAAATGATACGGCAACATACCAAGAACTTAC 204
Db 183 CTGGCTCAGAGAAACGAGCGCAGCGAAAGTATCCGAAACATATGCCGAGACATAT 242
QY 205 GGGCAAGACCTTCTCAAGACTTTGAACAAGAGCTCTTAACGATTTTCGAGAGCTATC 264
Db 243 GGGGAAGATCTTTCAGAAAGCGCTTGACAGAGAACTCTTAAGTACTTTGAGAGATCTG 302
QY 265 TTGTTGTGACTTTGAACCCGGTGAAGCTGATGCTTTATTTGGCTATAGAGCTACAAA 324
Db 303 CTCTGTGGGACCTCTGGATCTCGGAGCGGTATGCAATCTTGTCCATGAAAGCTACCAAG 362
QY 325 AGATGACCTTCAAGCAACCAAGTCTTATGGAAGTGTCTTGCAAGAGCATCAAGCGAG 384
Db 363 AGATGACTTTCAGCACTGGGTTCTCATGAAATGCTTGACAGAGCTTCAATAGAG 422
QY 385 CTGCTTCAAGCTAGGCAAGCTTACCAATGCTCGCTACAGAAAGTCTCTTGAGAGAGCTT 444
Db 423 TTATTCATGTTGAGGAGCGCTTATCATGCTCGTTATAGAAATCTCTTGAGAGAGCATC 482
QY 445 GCTCACCACTACCGGTGACTTCAAGAAAGCTTTGTTTCTTTGTTACCTCATACAG 504
Db 483 GCATATCACACTACCTGGGATTTCCGCAAGCTCTGTTCTCTGGCAAGTACTTTCCG 542
QY 505 TAGAAGAGAGATGAAGTGAACATGACATTTGGCTTAAGCAAGAACTAAGCTGCTCATGAG 564
Db 543 TATGAGGGGCTTGAAGTGAACATGACATTTGGCGAGATCAGAGGCTTAAGATCTTCATGAG 602
QY 565 AAAATCAAGAGCAAGCACTCAATGATGAGATGTTATTTAGATCTTTGTCACAAAGAGC 624
Db 603 AAGATTCACAGAAAGGCTTATCATATCATGATGAGCTCATCGAATTTGTTACTACAAAGAT 662
QY 625 AAAAGCTCAGTCAATGCTACTTTTAAACGTTTACCAAGATGATATGCGCAGAAATCTTC 684
Db 663 AAAAGCTCAGTCAATGCTACTTTTAAACGTTTACTCAACAGATGATTTGGGAATGCCATCAAC 722
QY 685 AAAAGCTTGAAGAGAGAGATGATGATGACATGCTTTCGACTTTTGAAGTCAACAT 744
Db 723 AAGGATCT--GAAAGCTGATCCAAATGATGATTTCTGAATGCTGAGATCAGCAAT 779
QY 745 CAGTGTCTGACAAAGCCAGAGCTTTTACTTTGATGATGCTTCTGTTCAAGCAATCAAAA 804
Db 780 AAGTGTGATCTTATTCAGAGAGTACTTTGAAGAGTATGCGTCTGCGCATCAACAG 839
QY 805 ACTGGAATGATGAGAGCACTCATGATTTGTCACAGAGAGCTGAGATTTGACTTG 864
Db 840 CTGGGAACAGATGAATGGGCTCTTTCACAGAGTATGCACTGCTGGGGAAGTTGACATG 899
QY 865 AAGGTCTTGAAGAGAGTACAGCGCAGAGACAGCAATCTTTGAGAAAGCTATTTAC 924
Db 900 CACCGGATCAAAAGAACATGACCAAAAGAGACAGTCTTCTTAATCGGGCTATTTGCA 959
QY 925 AAAAGCACTGCGAGATTTACAGAGAGATGCTGTCGCACTTCTCGGTAAAGTATGCT 984
Db 960 AATGACACATCTGGGGACTACAGAGAGATGCTTCTGCTTTGGTTGAGACATGAGATGCT 1019
QY 985 TAA 987
Db 1020 TGA 1022
```

```
RESULT 8
US-10-393-840-28
; Sequence 28, Application US/10393840
; Publication No. US2003022922A1
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1012c3
; CURRENT APPLICATION NUMBER: US/10/393,840
```

```
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 2588
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-393-840-28

Query Match 38.4%; Score 471.8; DB 17; Length 2588;
Best Local Similarity 68.6%; Pred. No. 8.7e-123;
Matches 666; Conservative 0; Mismatches 302; Indels 3; Gaps 1;

QY 30 AAAAATGGGACTCTTAAGGTTTCTGATTCGTCTGCTCTTCTGATGATGCTGAGCA 89
Db 1386 AACAAATGCTGACTTCACCGTCCGACGACCTGCCCTTGAAGCGATGCTGAGCA 1445
QY 90 ATTGGAACCGCTTTTGAAGATGGGTTACGAACGAGACTTATCATCTTGGC 149
Db 1446 GCTCGGACAGCTTTCGAGAGATGGGAAACAAATGAGAGCTATCATCTATTTGG 1505
QY 150 TCACAGAAATGCTGAACAGAGAAATCATCAGGCAAGATACCAAGAACTTACGCGCA 209
Db 1506 TCATAGGAATGCGGGCAGAGAGAGCTATTTGGCAACCTATGCGAGACTTACGCGCA 1565
QY 210 AGACCTTTCAGAGACTTTCGACAGAGAGCTCTTAACATTTTCGAGAGCTATCTTGT 269
Db 1566 GGAACCTCTCAAGGATTTGACAGAGAACTTACCAATGATTTTCGAGAGGCTGTGCT 1625
QY 270 GTGGAATCTTGAACCCGAGTGAAGTATGCTTATTTGGCTATTTGAAGTACAAAAGATG 329
Db 1626 TTGATCTCTTGAATCCGGCTGAGAGCTGATGCTGATCTTTGGAGATGAGACGAAAAGATG 1685
QY 330 GACTTCAGCAACCAAGTTCTTATGGAATGCTTTCGACAGAGACATCAACGAGCTGCT 389
Db 1686 GACTTCAGCAACCAAGTTCTTATGGAATGCTTTCGACAGAGCTTCCGCAAGTGTGCT 1745
QY 390 TCAAGTGGGCAAGCTTACATGCTCCGTAACAAGTCTCTTGAAGAGAGTGTGCTCA 449
Db 1746 TATGCAAGACAGATATCATGATCCGATACAAAGTCTGAGAGAGAGCTGCTGCTCA 1805
QY 450 CCACACTACCGGTGACTTTCAGAAAGCTTTGTTTCTCTTTTACCTCATACAGGTACGA 509
Db 1806 CCACACTACGAGATTTTCGTAAGTGTGCTGCTCTCTTGTGAGCTCTTACCGTTATGA 1865
QY 510 AGGAGATGAAGTGAACATGACATTTGGCTTGAAGAGAGTAAAGCTGCTCATGAGAAAT 569
Db 1866 TGAAGATGAGTGAATATATGACTTTGGCAAAACAGAGGCTTAAGATATCTCAAGAAAGAT 1925
QY 570 CAAGACAGACACTACATGATGAGAGATTTTGAATCTTGTCCACAAAGAGCAAGC 629
Db 1926 CTCAAGAGAGGCTTATGCGCAATGAGATCTCAAGAGATTTTGTACTAGAGCAAGC 1985
QY 630 TCAGATCAATGCTACTTTTAAACGTTTACCAAGATGATCATGCGAGAAATCTCAAGAG 689
Db 1986 AAGGTCAATGCTTACGCTGATATCATCAAAATAGATTTGAAATGATATCAACAGGA 2045
QY 690 TCTTGAAGAGAGATGATGATGACAGAGTCTTGTGACTTTTGAAGTCAACATTCAGTG 749
Db 2046 TTT--GAAATCTGATCCAAAGAGCGGTTCTTACTATATACGAGCTACAGTAAAGTG 2102
QY 750 CTTCGACAGACAGAGCTTACTTTGATGATTTCTTGTGATGACAAATTAACAAATCGG 809
Db 2103 CCTGACTGCGCTTGAAGAGATTTTGAAGAGGTTCTTGTCTTACGATCAATTAACGAGG 2162
```

Qy	810	AACATGATGAAGGAGACATCTAGAAATTGTGACCAACAAGCTGAGATTGACTTGAAGGT	866
Db	2163	AACAGATGAAGGGGCTCTGACCCAGAGTGTGTCTACCAAGGCTCGAGGTGACATGAAGTT	2222
Qy	870	CATTGAGAGAGATACCAAGCCGAGGAACAGCATTTCTTTGGAGAAGCTATTACCAAGA	929
Db	2223	TATAGTAGGAGGATACCAAGAGAGGAATATGATTCCTCTCGATTCGTGCATTGTCAAGA	2288
Qy	930	CACCTGCTGAGATTACGAGAAAGATGCTCGTGCACCTTCTCGGTGAAGATGATGCTTAAATC	989
Db	2283	CACCTACTGAGACATATGAAAAAATGCTTCTGCGCATTTGATGGCCACGTGAGGCTTGATT	2342
Qy	990	AATCAATCCTC 1000	
Db	2343	TACAACTACTC 2353	
RESULT 9			
US-09-770-961-765/c			
; Sequence 765, Application US/09770961			
; Publication No. US20030115639A1			
; GENERAL INFORMATION:			
; APPLICANT: Goriach, Jörn			
; APPLICANT: An, Yong-Qiang			
; APPLICANT: Hamilton, Carol M.			
; APPLICANT: Price, Jennifer L.			
; APPLICANT: Raines, Tracy M.			
; APPLICANT: Yu, Yang			
; APPLICANT: Rameaka, Joshua G.			
; APPLICANT: Page, Amy			
; APPLICANT: Matthew, Abraham V.			
; APPLICANT: Ledford, Brooke L.			
; APPLICANT: Moesner, Jeffrey P.			
; APPLICANT: Haas, William David			
; APPLICANT: Garcia, Carlos A.			
; APPLICANT: Krickler, Maja			
; APPLICANT: Slader, Ted			
; APPLICANT: Davis, Keith R.			
; APPLICANT: Allen, Keith			
; APPLICANT: Hoffman, Neil			
; APPLICANT: Hurban, Patrick			
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis			
; TITLE OF INVENTION: thaliana			
; FILE REFERENCE: 2026 (PARA-01SPRV)			
; CURRENT APPLICATION NUMBER: US/09/770,961			
; CURRENT FILING DATE: 2001-01-26			
; PRIOR APPLICATION NUMBER: 60/178,466			
; PRIOR FILING DATE: 2000-01-27			
; NUMBER OF SEQ ID NOS: 999			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 765			
; LENGTH: 487			
; TYPE: DNA			
; ORGANISM: Arabidopsis thaliana			
US-09-770-961-765			
Query Match 38.2%; Score 470; DB 10; Length 487;			
Best Local Similarity 100.0%; Pred. No. 1,1e-122;			
Matches 470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	761	CAGAGCTTTACTTTGTGCGATGTCTTCTGTTGACGAAATCAACAAAACTGAACTGATGAG	820
Db	487	CAGAGCTTTACTTTGTGCGATGTCTTCTGTTGACGAAATCAACAAAACTGAACTGATGAG	428
Qy	821	GAGACCTACTAGAAATTGTGACCAACAAGCTGAGATGACTTGAAGCTCAATTGGAAGG	880
Db	427	GAGACCTACTAGAAATTGTGACCAACAAGCTGAGATGACTTGAAGCTCAATTGGAAGG	368
Qy	881	AGTACCAAGCGAGAAACAGCATTTCTTTTGAAGAAAGCTATTACCAAAAGACACTCGTGAG	940
Db	367	AGTACCAAGCGAGAAACAGCATTTCTTTTGAAGAAAGCTATTACCAAAAGACACTCGTGAG	308
Qy	941	ATTACAGAGATGCTCTCGCGCACTTCTCGGTGAAGATGATGCTTAATCAATCAATCTCTC	10000

```

Db      | ATTAGAGAAAGATGCTCGTGGCATTCTTCGGTGGAAGATGATCTTAATCAATCAATCCCTC 248
| |||||
Qy      | 1001 CACAGAGAAAATAAGCTGTCTCTACAGCTTCGTATATCTCATATCTCCCTCTCTCTCT 1060
| |||||
Db      | 247 CACAGAGAAACAATAAGCTGTCTCTACAGCTTCGTATATCTCATATCTCCCTCTCTCTCT 188
| |||||
Qy      | 1061 TTGATGAGTTCCAATTCGTTGATTTTGTTCCTAACAAAAAACCCTGTTGTTTCTGTGTG 1120
| |||||
Db      | 187 TTAGATGACTTCAAATTCGTTGATTTTGTTCCTAACAAAAAACCCTGTTGTTTCTGTGTG 128
| |||||
Qy      | 1121 TGTTTSGAGTCCCTTAATAAGSCAAAGAAGAGACAGAGAAACGATGCGTCTTA 1180
| |||||
Db      | 127 TGTTTSGAGTCCCTTAATAAGSCAAAGAAGAGAGACAGAGAAACGATGCGTCTTA 68
| |||||
Qy      | 1181 AGTTATATATATATGAAGAGCATTTGCCCTTAAAAAAAAAAAAAAAAAAAAAA 1230
| |||||
Db      | 67 AGTTATATATATATGAAGAGCATTTGCCCTTAAAAAAAAAAAAAAAAAAAAAA 18
| |||||

RESULT 10
US-10-219-220-17
; Sequence 17, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Plim, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of plant development
; FILE REFERENCE: 11000.1022c1
; CURRENT APPLICATION NUMBER: US/10/219,220
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-219-220-17

Query Match          38.1%; Score 468.6; DB 14; Length 1293;
Best Local Similarity 68.4%; Pred. No. 4.8e-122;
Matches 664; Conservative 0; Mismatches 304; Indels 3; Gaps 1;

Qy      | 30 AAAAAAGCGACTCTTAAGTCTTCTGATTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 69
| |||||
Db      | 91 AACCAATGTCGACTCTCACCGTCCGCACCACTGCCCTGTAGCCGATGACTCGAGCA 150
| |||||
Qy      | 90 ATTGAGAACCCCTTTTGAAGATGGGGGTACGAAACGAGACTTGATCATATCATCTTGGC 149
| |||||
Db      | 151 GCTCCGAGACGCTTCGAGAGATGGGAAACAAATGAAAGCTGATCATATCATATTTGG 210
| |||||
Qy      | 150 TCACAGAAAGTCTGAACAGAGAAAGTCATCAGGCAAGCATACACAGAAACCTACGGCA 209
| |||||
Db      | 211 TCATAGGAATCGGCGCAGAGAAAGCTGATTCGGCAACCTATGCGAGACTTACGGCGA 270
| |||||
Qy      | 210 AGACCTTCTCAAGACTCTTGACAGAGAGCTCTCTACGATTTTCAAGAGAGCTATCTTGT 269
| |||||
Db      | 271 GGACCTCTCAAGGCAATGGAACAGAACTTAACCAATATTTTCGAGAGGCTGTGGTCT 330
| |||||
Qy      | 270 GTGAGCTCTTAACCCGGTGAGCGTGAATGCTTTATTTGGCTAATGAAGCTACAAAAAGATG 329
| |||||
Db      | 331 TTGATCACTTATCCGCTGAACGTGATGCTACTTGGCGAATGAAGCAGCAAAAAGATG 390
| |||||
Qy      | 330 GACTTCAAGCAACCAAGTTCTTATGAAAGTTGCTTGCAACAGACATCAACGAGCTGCT 389
| |||||
Db      | 391 GACTTCAAGCAACCAAGTTCTCATGAAATAGCTGCGACGAGGTCTCCGACAGCAATGGCT 450
| |||||
Qy      | 390 TCAAGCTAAGCAACTTACATGCTGCTGCTACAAAGATGCTCTTGAAGAGAGAGTTGCTCA 449
| |||||
Db      | 451 TAGGCAAGAACACATATCATATGCCCCGATTAACGAAGATCAATGGAAGAGAGCTGCTCA 510
| |||||

```

```

QY 450 CCACGTACCGGAGCTTCAGAAAGCTTTGGTTTCTCTTGTACCTCATACAGTACGA 509
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 511 CCACAAATCGAGATTTTCGTAGTGTGCTGTAACCTCTTGGAGCTCTTACCTAATGA 570
QY 510 AGGAGATGAAGTAAATGATGATGCTTAAAGCAAGAGTAACTGTCATGAGAAAT 569
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 571 TGGAGATGAGTAAATATGATCTTGGCAAAAGAGAGGCTAAGATACCTCAAGAAAGAT 630
QY 570 CAAGCAAGAGCTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 629
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 CTCAGAGAGGCTTATGCGCATGAGATCTCATAGAGATTTTGGCTACAGAGCAAAAC 690
QY 630 TCAGATCAATGCTACTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 689
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 691 ACAGGTCAATGCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 750
QY 690 TCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 749
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 751 TTT---GAAAACTGATCCAAAAAGACGGTTCTTACTATAGAGCTACAGTAAAGTG 807
QY 750 CTGGAACAAGACCGAGCTTACTTTGTGATGTTCTTGTGCTTGTGCTTGTGCTTGTGCT 809
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 808 CCGACTCCCGCTGAGAAAGATTTTGAAGGTTCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 867
QY 810 AACTGATGAAGAGAGCACTCACTGAATGTGACCAAGAGCTGATGATGATGATGATGAT 869
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 868 AACAGATGAAGGGGCTCTGACAGAGTATGTTGCTACAGAGCCGAGAGTGAATGAAGTT 927
QY 870 CATTTGAGAGAGAGTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 929
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 928 TATTAATGAGAGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 987
QY 930 CACTCGTGAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 989
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 988 CACTACTGAGAGACTAAGAAAAATGCTTCTGCAATGATTTGCGACGTGAGGCTTGATT 1047
QY 990 AATCAATCTC 1000
    |||||
Db 1048 TACAAGTACTC 1058

```

```

RESULT 11
US-10-393-840-109
; Sequence 109, Application US/10393840
; Publication No. US2003022922A1
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.101263
; CURRENT APPLICATION NUMBER: US/10/393,840
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-393-840-109

```

```

Query Match      38.1%; Score 468.6; DB 17; Length 1293;
Best Local Similarity 68.4%; Pred. No. 4.8e-122;
Matches 664; Conservative 0; Mismatches 304; Indels 3; Gaps 1;

```

```

QY 30 AAAAAAGGAGCTTTAAGTTTCTGATTTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 91 AACAAATGCTGATCTTCAACCGTCCGACAGCACTGCCCCCTGATGCGATGATGAGAGA 150
QY 90 ATTTGAACCCGCTTTTGAAGATGGGGTACCAAGAGAGATTTGATCATATCAATCTTGGC 149
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 151 GCTCCGAGAGAGCTTCCGAGATGGGGAAACAATAGAGAGCTGATCATATCATATTTGGG 210
QY 150 TCACAGAAATGCTGAACAGAGAAAGTATCATGAGGAGAGATACCAAGAACTTACGGCGA 209
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 TATAGAAATGGGGGCGAGAGAGAGTATGCGAAACCTATGCGAGACTTACGGCGA 270
QY 210 AGACCTTCTCAAGACTTTTGAACAAGAGCTCTTCAAGATTTTGAAGAGATATCTTGT 269
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 GAGACCTCTCAAGAGATTTGAACAAGAGCTTCAAGATTTTGAAGAGCTGATGATGAT 330
QY 270 GTGACCTTTGAACCCGCTTGAAGAGCTTATTTGCTTAAATGAAGTCAAAAAAGATG 329
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 TTGCTCACTGATCCGGCTGAACGTGATGCGTACTTGGCGAATGAAGGAGCAAAAAAGATG 390
QY 330 GACTTCAAGCAACCAAGTTCTTATGAAAGTTGCTTGAACAAGGATCATCAAGCAATCAAGCTGCT 389
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 GACTTCAAGCAACCAAGTTCTTATGAAATGCTTGAAGAGCTTCCGAGAGATGCT 450
QY 390 TCACGCTAGGCAAGCTTACCATGCTGCTACAAAGAGCTCTTGAAGAGAGCTTCTCA 449
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 TATGCAAGAGCAAGGATATCATGCCCCGATACAAAGAGTCAATGAAGAGAGAGCTGCTCA 510
QY 450 CCACACTACCGGTGCTTCAAAAGCTTTTGTCTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGT 509
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 511 CCACAACTGAGAGATTTTGTGATGCTGCTTGTCTTGTGAGCTCTTCAAGCTTCAAGTGA 570
QY 510 AGGAGATGAAGTGAACATGACATTTGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 569
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 571 TGAAGATGAGTGAATATGATCTTTGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 630
QY 570 CAAGCAAGAGCACTCAATGATGAGAGATGTTATGAAATCTTGTCCAAAGAGAGAGAG 629
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 CTCAGAGAGAGGCTTATGAGCATGAGATCTCATAGAGATTTTGGCTACTAGAGAGAGAG 690
QY 630 TCAGATCAATGCTACTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 689
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 691 ACAGGTCAATGCTTACGCTGAATCACTACAAAAATGATTTGAAATGATATCAACAGAG 750
QY 690 TCTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 749
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 751 TTT---GAAAACTGATCCAAAAAGACGGTTCTTACTATAGAGCTACAGTAAAGTG 807
QY 750 CTGGAACAAGACCGAGCTTACTTTGTGATGTTCTTGTGCTTGTGCTTGTGCTTGTGCT 809
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 808 CCGACTCCCGCTGAGAAAGATTTTGAAGGTTCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 867
QY 810 AACTGATGAAGAGAGCACTCACTGAATGTGACCAAGAGCTGATGATGATGATGATGAT 869
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 868 AACAGATGAAGGGGCTCTGACAGAGTATGTTGCTACAGAGCCGAGGTTGAGATGAAGTT 927
QY 870 CATTTGAGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 929
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 928 TATTAATGAGAGATACAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 987
QY 930 CACTCGTGAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 989
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 988 CACTACTGAGAGACTAAGAAAAATGCTTCTGCAATGATTTGCGACGTGAGGCTTGATT 1047
QY 990 AATCAATCTC 1000
    |||||
Db 1048 TACAAGTACTC 1058

```

```

RESULT 12
US-10-425-114-10772
; Sequence 10772, Application US/10425114
; Publication No. US2004003488A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ. ID NOS: 73128
; SEQ. ID NO 10772
; LENGTH: 1145
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700970311_FLI
US-10-425-114-10772

```

```

Query Match      36.6%; Score 450.4; DB 18; Length 1145;
Best Local Similarity 67.8%; Pred. No. 6.6e-117;
Matches 660; Conservative 0; Mismatches 311; Indels 3; Gaps 2;

```

```

QY 14 ACACTAAAGTAAAGAAAATGCGACCTTAAAGTTTCTGATCTCTCTCTCT 73
DB 34 AAACAAAACAGAAACACATGCGACCTTAAAGTTCTCTGACCTCTCTCTCT 93
QY 74 CTGATGATGCTGACATTTGAGAACCGCTTTTGAAGATGGGTAACAGAGACTTGA 133
DB 94 CAATGATTTGTGAGAGCTCCGAAAAGCTTCTCAGTTTGGGAACTTAAGAGAGTGA 153
QY 134 TCATATCATCTTGGCTCACAAGAGTCTGAACAGAGAAAAGTATCAGGCAACATACC 193
DB 154 TTGTATGATCTTGGCTCATAGAAATGCTCTCAGAGAACTTATCAGAAACTTATG 213
QY 194 ACGAAGCTACGCGGAGAGACTTCTCAAGACTCTTGAAGAGAGCTCTCAAGATTTG 253
DB 214 CCGAGACTTAATGAGAAAGATCTCTCAAGGCTTGAAGAAAGTCTACAGATGATTTG 273
QY 254 AGAGAGCTATCTTGGTGTGAGACTTGTGAACCGGAGAGCTGATGCTTATGAGTATG 313
DB 274 AGAGGCTGTTCTATCTTTGACACTTATCTGTGACAGTATGCAATTTTGGCAACG 333
QY 314 AAGCTACAAAAGATGAGACTTCAAGCAACCAAGTTCTTATGAGAGTCTTGAACAAGA 373
DB 334 AGCGACTTAATAATGAGACTTCAAGCAATCAGGTTCTGTGAAATAGCTGCACTAGGT 393
QY 374 CATCAACGCGCTGCTTCAAGCTTGAAGCAAGCTTACATGCTGCTTACAGAAAGTCTTGG 433
DB 394 CCTCTGAACAAATGTTGTGCTGCGAGAGAGCTTACCAATGTTCTTATTAAGAGTCTCTGG 453
QY 434 AAGAGAGCTTGTGCTCAACACTACCGGATGCTTCAAGAAAGCTTTGTTGTTCTCTGTTA 493
DB 454 AGAGGATGTTGCTCATCAACAACAGAGACTTCCGTAAGCTCATACCTCTGTTA 513
QY 494 CCTCATACAGTACGAAGAGATGAACTGAATGACATTTGCTTGAAGCAAGAGCTTAAGC 553
DB 514 GTTCTTATGATTAAGAGAGATGAGGTCAACTTGAACCTT-GCAAAAACAGCAAAAAT 572
QY 554 TGTTCATGAGAAATCAAGAGACATCAATGATGAGATGTTATTAAGATCTTGT 613
DB 573 TGTCTGATGAGAAATTTCAACAGAGCTTATTAATGACAGCTTCAATCAGATTTTGG 632
QY 614 CCAACAAGAAAGCTCAGATCAATGCTTATTAACGTTTCAAGATGATCATGGC 673
DB 633 CCAACAAGAGACAGACAGATTAATGCTTCTTTGAATCCTACCAAGATGATTTGG 692
QY 674 AGGAAATTTCTCAAGAGCTTGAAGAAAGAGATGATGACAAAGTCTCTTCACTTTTGA 733
DB 693 ACAAGATATCAACAAGAGACTGA--AGGCTGATCCAAAGAGAGAGTCTCTTACTTA 750

```

```

QY 734 GGTCAACATTCAGTGTCTTGAACAAGACAGAGCTTACTTTGTCATGTTCTTCTGTCAG 793
DB 751 GAGCAACTGTGATGATCTTGAATCCCTCCGAGAAAGTACTTCAAGAGTTGTTCCGTTGG 810
QY 794 CAATCAACAAACTGGAAGTGAAGAGACATCTAGAAATTTGACACAAAGACTG 853
DB 811 CGATCAACAAGAGAGAACGATGAGAGAGCCCTTCAAGAGTGTGGCAACAAGGCTG 870
QY 854 AGATTGACTTGAAGTCTTGAAGAGAGATCCAGCGCAAGAACAGATCTCTTGGAGA 913
DB 871 AGGTTATTTGAAAAAATATGACAGATGATATCAAAAGAGAGCAAGTGTCTCTTGAAC 930
QY 914 AGCTATTACCAAGACATCTGAGATTAAGAGAAAGATGCTCGGACTCTCGGTC 973
DB 931 GTGCTATTTGTTAAGGACACACCGGTGACTATGAGAAAATGCTTGTGGGCTTTTAGAC 990
QY 974 AAGATGATGCTTAA 987
DB 991 ATGATGATGCTTGA 1004

```

## RESULT 13

```

US-10-424-599-70457
; Sequence 70457, Application US/10424599
; Publication No. US20040031072A1

```

```

; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ. ID NOS: 285684
; SEQ. ID NO 70457
; LENGTH: 1321
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34633C.1
US-10-424-599-70457

```

```

Query Match      35.0%; Score 430.8; DB 18; Length 1321;
Best Local Similarity 66.9%; Pred. No. 2.7e-111;
Matches 644; Conservative 0; Mismatches 312; Indels 6; Gaps 2;

```

```

QY 26 GAAGAAAATGCGGACTCTTAAGTTTCTGATTTCTGTGCTCTCTCTGATGATGCTG 85
DB 82 GATCGAAAATGCGGAGCTTGAAGGTTCTGCACAACTTCTTCTCATTTGAAAGACAGTG 141
QY 86 AGCAATTTGAGAACCGCTTTTGAAGATGGGCTACGAAACAGAGACTTATCATCAATCT 145
DB 142 AGCAATTTGAGAAAAGCTTTTCAAGGCTGGGAAACGAATTAAGATTTGATTAATCAATC 201
QY 146 TGGCTCAAGAAAGTGTGAAAGAGAGAAAGTATCATCAGGACATTAACAGAACTTACG 205
DB 202 TGGGTCAAGAAATGGGCTCAGCGTAAGTATGTTGTAAGCTTATTAACCAACCCATG 261
QY 206 GCGAAGACTTGTCAAGACTCTTGAACAAGAGCTCTTAACGATTTCCGAGAGACTATCT 265
DB 262 GAGAAAGATTTCTTAAAGATTTGAACAAAGAACTTTCAAGTACCTTGAAGCGGCTGTGT 321
QY 266 TGTGTGAGACTCTTGAACCGGCTGAGCGTATGCTTTATTTGGCTAATGAAGCTACAAAA 325
DB 322 TGTGTTGACATTTGATCTCTGAGCGTATGCTCTTTTACTATTAAGCACTTAAGA 381
QY 326 GATGACCTTCAAGCAACCAAGTCTTATTAAGAAAGTGTGTCACAAAGACATCAAGCAGC 385
DB 382 TGTGACTTCAAAACAATTTGGTCAATTTTGAATTTGCTTCACTAGATCTTCACTTGAATC 441
QY 386 TGTCTCAGCGTACGCAAGCTTACATGCTCGCTACAAAGAAAGTCTTGAAGAGAGCTTG 445

```



```
Db 442 TCCTTAGGAAAGCAAGCATATCAAGCCGTTTCAAAAAGTCCCTTGAAGAAGGTTG 501
Qy 446 CTCACACACTACCGGTGACTTACAAAGCTTTGGTTTCTCTTTACTCATACAGGT 505
Db 502 CCTATCATCTAAAGGGGACATCCGCAAGCTATTACTTCTCTTGGACATATTCGGT 561
Qy 506 ACGAAGAGA-TSAGTGAACATGACATTTGGCTTACCAAGCTAAGCTGGTCCATGAG 564
Db 562 ATGAGGGGATTTAGGTGAACATGACATTTGGCAAAATCTGAGGCTTAACTGCTTCAAG 621
Qy 565 AAAATCAAGACAAGACACTACATGATGAGATGTTATTAGAACTTTGCCAACAAGAC 624
Db 622 AAGATTGACAGAAAGGCTTACATGATGAGAGCTGATCAGGATTTTGTCAACAAGAGT 681
Qy 625 AAGCTCAGATCATGCTACTTTTAACTGTTACCAAGATGATATGCGGCAAAATTCG 684
Db 682 AAAGCAGATTACACAACAACATTAATCATGATCATGAGTTCCGGAATGCCATMAAC 741
Qy 685 AAGAGCTTGAGGAGAGATGATGATGACAAAGTTCCTTGCACTTTGA----GGTCAA 739
Db 742 AAGGATCTGAAAAATTGAACTCAAAAAGATGGAATATCTGCAATTTATTGAGGCAAG 801
Qy 740 CCATTGAGCTTGACAAGACAGAGCTTACTTTGCGATGTTCTTTCGTTACAGCAATCA 799
Db 802 CCATTGAGCTTGACTTACCTGAGAAATATTTGCAAAAGCTTGAAGGCTGGCTATAA 861
Qy 800 ACAAACCTGCACTGATGAAAGGACACTCACTGAAATTTGTGACACAAAGCTGAGATTG 859
Db 862 ACAAGCTAGGACTGATGAAAGGGCCCTTACTAGAGTGGTGAACAGAGCTGAGATTG 921
Qy 860 ACTTGAAGCTCATTTGAGAGGAGTACAGGCGCAGAAACAGATTCCTTTGAGAAAGCTA 919
Db 922 ATTGCAAGCAATTTGCGAAATCCAAAGAAACAGCATTTCTTGGACCGGCA 981
Qy 920 TTACCAAGACTCGTGAAGATTACGAGAAAGTCTGCGCACTTCTCGTGAAGATG 979
Db 982 TACCGAGTGACACTTTCGAGGATATCAAGAGACTTCTCTGCACTGTTGACATGAG 1041
Qy 980 AT 981
Db 1042 AT 1043

RESULT 14
US-10-219-220-20
; Sequence 20, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Iasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE REFERENCE: 11000.1022c1
; CURRENT APPLICATION NUMBER: US/10/219,220
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932
; NUMBER OF SEQ ID NOS: 250
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-219-220-20

Query Match 29.6%; Score 364.2; DB 14; Length 789;
Best Local Similarity 69.7%; Pred. No. 1,7e-92;
Matches 508; Conservative 0; Mismatches 218; Indels 3; Gaps 1;
```

```
Qy 85 GAGCAATTGAAACCGCTTTTGAAGATGGGTACAAAGAGACTTGATCATATCAATC 144
Db 124 GAGCAGCTCCAAAAAGCTTTTGCAGAGATGGGGAGCAATGAAATCTGATCATATCATA 183
Qy 145 TTGGCTCA CAGAAGTCTGAACAGAGAAAGTCACTACAGCAAGCATACAGAAACCTAC 204
Db 184 CTGCTTCACAGAAACGACGCGAGGAAAGTATCCACAACATATAGCCGACATAT 243
Qy 205 GCGCAAGACTTCTCAAGCTTTTGAAGAGAGCTCTTAACGATTTTGAAGAGACTATC 264
Db 244 GGGGAAGATCTTCTAAAGCCGTTGACAGAACTCTTAAGTACTTTGAGAGACTCTGTG 303
Qy 265 TTGTTGAGACTCTTGAACCCGTTGAGGATGATGCTTATTGGCTTAATGAAGCTCAAAA 324
Db 304 CTTCGTGAGACCTTGATCTGCGAGGCTGATGATCTTGTCCAAATGAAGCTACCAAG 363
Qy 325 AGATGACTTCAAGCAACCAAGTTCTTATGAAAGTGTCTTGCAAGAGACATCAACGAG 384
Db 364 AGATTGACTTGAGCAACGAGGTTCTCATGAAATTTGCTTGCAAGAGGCTTCAATGAG 423
Qy 385 CTGCTTACGCTGAGCAAGCTTACATGCTGCTACAAAGATCTTGAAGAGACGTT 444
Db 424 TTATTCATGTGAGCAGGCTTATCATGCTGTTATTAAGAAATCTTGAAGAAAGACATC 483
Qy 445 GCTCAACACTACCGGTGACTTGAGAAAGCTTTGTTTCTTGTACCTCATACAG 504
Db 484 GCATTCACACTACTGAGGATTTCCGCAAGCTGCTTGTCTTCTGCAAGTACCTTTGG 543
Qy 505 TACGAAGAGATGAAGTGAACATGACATTTGGCTTGAAGAAAGCTAAGCTGCTCATGAG 564
Db 544 TATGAGGGGCTGAGGTGAACATGACATTTGGGAGATCAGAGGCTTAAGATCTTATGAG 603
Qy 565 AAAATCAAGACAGACACTCAATGATGAGATGTTATTAGAAATTTGTCCACAAGAGC 624
Db 604 AAGATTCAGAAAGGCTTACATCATGATGATGATCAGAAATTTGATCAAGAAAGT 663
Qy 625 AAGGCTGAGATCATGCTTACTTTTAAACGTTTACCAAGATGATCATGGCAGAGAAATCTC 684
Db 664 AAGCTCAGCTTAAATGCAACCTCAATTAATCAATGATGATGATTTGGGAATGCCATCAAC 723
Qy 685 AAGAGCTTGAGAAAGAGATGATGATGACAAAGTTCCTTGCACTTTGAGAGTCAACAT 744
Db 724 AAGGATCT--GAGGCTGATCAATGATGATGATTTGAAACTGCTGAGATCAACAAAT 780
Qy 745 CAGTCTTG 753
Db 781 AAGTCTTG 789

RESULT 15
US-10-393-840-26
; Sequence 26, Application US/10393840
; Publication No. US20030229922A1
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; TITLE OF INVENTION: Modification of Plant Cell Wall Polysaccharides
; FILE REFERENCE: 11000.1012c3
; CURRENT APPLICATION NUMBER: US/10/393,840
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 789
```



TYPE: DNA  
ORGANISM: Eucalyptus grandis  
US-10-393-840-26

Query Match 29.6%; Score 364.2; DB 17; Length 789;  
Best Local Similarity 69.7%; Pred. No. 1.7e-92;  
Matches 508; Conservative 0; Mismatches 218; Indels 3; Gaps 1;

```
OY 25 AGAAGAAAATGGGACTCTTAAGGTTTCTGATTCCTGCTCCTTCTGATGATGCT 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 AGAAGAGAATGGGACTATGCGGTGCGACCTCGGTCCGCTCCGCTGAGGATGCC 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 85 GAGCAATTGAGACCGCTTTGAGGATGGGGTACGACGAGCTTGATCATATCATC 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 GAGCAGCTCCAAAAGCTTTGACAGATGGGGAGCAATGAGATCTGATCATATCCATA 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 145 TTGGCTCAGAGAGTGTGAACAGAGAAAGTCATCAGCAGAGCATACAGAACTTAC 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 CTGCTCAGAGAAAGCGAGCGAGAGTAATCCGACAAACATATGCCGAGACATAT 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 205 GGGCAAGACTTCTCAAGACTTTGACAGAGACTCTTAAAGATTTCGAGAGCTATC 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 GGGGAAGATCTTCTCAAGCGCTTACAGAGAACTCTAGTGACTTGGAGAGATCTGTG 303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 265 TTGTTGGACTCTTGAACCGGAGGAGGTGATGCTTATTGGCTAATGAAGCTACAAA 324
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 CTTCTGTGAGCCTGAGATCGCGAGGTGATGATCTTGTCAATGAAGCTACCAAG 363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 325 AGATGACTTCAGAGCAACAAGTTCTTATGGAAGTTGCTTGCACAGAGCATCAACGAG 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 364 AGATTGACTTGAAGCACTGGGTTCTCATGAAATGCTTGCAGAGATCTTCAATGAG 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 385 CTGCTTACGCTGAGCAAGCTTACATGCTGCTACAGAAAGTCTTTGAAGAGACGTT 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 TTATTCAATGAGGAGGAGGCTTATCATGCTGTTATGAAGATCTTTGAAGAGACATC 483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 445 GCTTACACACACTACGGTGAATTGAAAGCTTTGGTTCTCTTTACTCTATACAG 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 484 GCATATACACTACTGGGAGATTTCCGAGAGCTGCTTGTCTTGGCAAGTACCTTTGG 543
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 505 TACGAAGAGATGAAGTGAATGATGATGATGCTTAAGCAAGAGCTTAAGTGTCTATGAG 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 544 TATGAGGGGGCTGAGTGAACATGATGATGCTTAAGGCTTAAGATCTTCAATGAG 603
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 565 AAAATCAAGACAAAGCACTACATGATGAGAGTATTAAGATCTTGTCCACAAGAGC 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 604 AAGATTACGAGAGAGGCTTACATCATGATGAGCTCATCAGAAATTGTTACTACAAAGAT 663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 625 AAAGCTCAGATCAATGCTACTTTAACCGTTACCAAGATGATCATGCGAGAAATCTC 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 664 AAAGCTCAGCTTAATGCAACCTCAATTAACAACAATGAGTTGGAAATGCCATCAAC 723
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 685 AAGAGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 724 AAGGATCT--GAAGGCTGATCCAAATGATGATTTCTGAAGACTGCTGATGATGAT 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 745 CAGTCTTG 753
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 781 AAGTCTTG 789
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: August 22, 2005, 07:12:44  
Job time : 2743 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 20:23:59 / Search time 165 Seconds

(without alignments)  
743.049 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611  
Sequence: 1 MATLKVSQSVPAIPDDAEQL.....KDTGRDYKMLVALGDDDA 317

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_16Dec04:\*  
1: geneseq1980s:\*  
2: geneseq1990s:\*  
3: geneseq2000s:\*  
4: geneseq2001s:\*  
5: geneseq2002s:\*  
6: geneseq2003as:\*  
7: geneseq2003bs:\*  
8: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1611	100.0	317	3	AAG09200	Aag09200 Arabidops
2	1611	100.0	317	3	AAG08874	Aag08874 Arabidops
3	1611	100.0	317	8	ADU53872	Adj53872 Plant lip
4	1611	100.0	317	8	ADN74567	Adn74567 Thale cre
5	1611	100.0	327	3	AAG09199	Aag09199 Arabidops
6	1611	100.0	328	3	AAG08873	Aag08873 Arabidops
7	1185.5	73.6	316	3	AAB16294	Aab16294 Eucalyptu
8	1171.5	72.7	316	3	AAB16321	Aab16321 Pinus rad
9	1171.5	72.7	316	4	AAB55730	Aab55730 Annexin-1
10	1171.5	72.7	316	7	ADB94698	Adb94698 Programme
11	1132.5	70.3	316	7	ADB94896	Adb94896 Programme
12	1071	66.5	211	3	AAG09201	Aag09201 Arabidops
13	1071	66.5	211	3	AAG08875	Aag08875 Arabidops
14	1071	66.5	211	3	AAG06311	Aag06311 Arabidops
15	1067	66.2	317	3	AAG33493	Aag33493 Arabidops
16	1067	66.2	317	3	AAG34920	Aag34920 Arabidops
17	1067	66.2	317	8	ADN74791	Adn74791 Thale cre
18	1034.5	64.2	316	3	AAG27921	Aag27921 Arabidops
19	988	61.3	302	3	AAG49231	Aag49231 Arabidops
20	988	61.3	304	3	AAG49230	Aag49230 Arabidops
21	928.5	57.6	284	3	AAG27922	Aag27922 Arabidops
22	866.5	53.8	239	3	AAB16292	Aab16292 Eucalyptu
23	866.5	53.8	239	4	AAB16324	Aab16324 Pinus rad
24	866.5	53.8	239	4	AAB55733	Aab55733 Annexin-1
25	866.5	53.8	239	7	ADB94701	Adb94701 Programme

#### ALIGNMENTS

26	787	48.9	240	3	AAG49232	Aag49232 Arabidops
27	787	48.9	240	3	AAG23494	Aag23494 Arabidops
28	787	48.9	240	3	AAG49203	Aag49203 Arabidops
29	784.5	48.7	239	3	AAG27923	Aag27923 Arabidops
30	770	47.8	153	3	AAG06312	Aag06312 Arabidops
31	721	44.8	242	3	AAB16293	Aab16293 Pinus rad
32	712	44.2	220	3	AAG33495	Aag33495 Arabidops
33	712	44.2	220	3	AAG49204	Aag49204 Arabidops
34	625	38.8	204	3	AAG17531	Aag17531 Arabidops
35	625	38.8	206	3	AAG17530	Aag17530 Arabidops
36	602.5	37.4	321	3	AAG30142	Aag30142 Arabidops
37	602.5	37.4	321	3	AAG38352	Aag38352 Arabidops
38	602.5	37.4	345	3	AAG38351	Aag38351 Arabidops
39	573	35.6	184	3	AAB16325	Aab16325 Pinus rad
40	573	35.6	184	4	AAB55734	Aab55734 Annexin-1
41	573	35.6	184	7	ADB94702	Adb94702 Programme
42	563	34.9	184	3	AAB16290	Aab16290 Human hea
43	512	31.8	316	7	ADJ69605	Adj69605 Human hea
44	509	31.6	327	4	ABR10307	Abri10307 Human CDV
45	509	31.6	327	5	ABP66894	Abp66894 Human pol
ALIGNMENTS						
RESULT 1						
AAG09200	AAG09200 standard; protein; 317 AA.					
ID	AAG09200; (first entry)					
AC	AAG09200;					
XX	XX					
DT	17-OCT-2000 (first entry)					
XX	XX					
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 7039.					
XX	XX					
KW	Protein identification; signal transduction pathway; metabolic pathway;					
KW	hybridisation assay; genetic mapping; gene expression control; promoter;					
KW	termination sequence.					
OS	Arabidopsis thaliana.					
XX	XX					
PN	EP1033405-A2.					
XX	XX					
PD	06-SEP-2000.					
XX	XX					
PF	25-FEB-2000; 2000EP-00301439.					
XX	XX					
PR	25-FEB-1999;	99US-0121825P.				
PR	05-MAR-1999;	99US-0123180P.				
PR	09-MAR-1999;	99US-0123548P.				
PR	23-MAR-1999;	99US-0123788P.				
PR	25-MAR-1999;	99US-0126264P.				
PR	29-MAR-1999;	99US-0126785P.				
PR	01-APR-1999;	99US-0127462P.				
PR	06-APR-1999;	99US-0128234P.				
PR	08-APR-1999;	99US-0128714P.				
PR	16-APR-1999;	99US-0129845P.				
PR	19-APR-1999;	99US-0130077P.				
PR	21-APR-1999;	99US-0130449P.				
PR	23-APR-1999;	99US-0130510P.				
PR	28-APR-1999;	99US-0130891P.				
PR	30-APR-1999;	99US-0131449P.				
PR	30-APR-1999;	99US-0132048P.				
PR	04-MAY-1999;	99US-0132407P.				
PR	05-MAY-1999;	99US-0132484P.				
PR	06-MAY-1999;	99US-0132485P.				
PR	06-MAY-1999;	99US-0132486P.				
PR	07-MAY-1999;	99US-0132487P.				
PR	11-MAY-1999;	99US-0132566P.				
PR	14-MAY-1999;	99US-0134218P.				
PR	14-MAY-1999;	99US-0134219P.				
PR	14-MAY-1999;	99US-0134221P.				

PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135533P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-014187P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147329P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153707P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.

PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 1611; DB 3; Length 317;  
Best Local Similarity 100.0%; Pred. No. 1,2e-139;  
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLKVDSPVAPSDDAEQRTAFEGWGTEDELLISILARSAPQKRVIAQAHETYGED 60  
1 MATLKVDSPVAPSDDAEQRTAFEGWGTEDELLISILARSAPQKRVIAQAHETYGED 60  
Db 1 MATLKVDSPVAPSDDAEQRTAFEGWGTEDELLISILARSAPQKRVIAQAHETYGED 60  
QY 61 LKLTLDKELSNDFERRILLMTLEPERDALLANEATKRMTSNNOVLMEVACTRTSTOLH 120  
61 LKLTLDKELSNDFERRILLMTLEPERDALLANEATKRMTSNNOVLMEVACTRTSTOLH 120  
Db 61 LKLTLDKELSNDFERRILLMTLEPERDALLANEATKRMTSNNOVLMEVACTRTSTOLH 120  
QY 121 APOAYHARYKKSLEEDVAHHTTGDPRKLVLVSYRYEGDEVNMTLAKQBAKLVBKIK 180  
121 APOAYHARYKKSLEEDVAHHTTGDPRKLVLVSYRYEGDEVNMTLAKQBAKLVBKIK 180  
Db 121 APOAYHARYKKSLEEDVAHHTTGDPRKLVLVSYRYEGDEVNMTLAKQBAKLVBKIK 180  
QY 181 DKXINDVDVIRILSTRKQAQINATFNRYODDHGEELIKSLBEGDDDKFLALLRSTIOCL 240  
181 DKXINDVDVIRILSTRKQAQINATFNRYODDHGEELIKSLBEGDDDKFLALLRSTIOCL 240  
Db 181 DKXINDVDVIRILSTRKQAQINATFNRYODDHGEELIKSLBEGDDDKFLALLRSTIOCL 240  
QY 241 TRPELVYVDVLRSAINTGTDEGALTRIIVTTBAIRDKLVIGEYORNSIPILEKAITKDT 300  
241 TRPELVYVDVLRSAINTGTDEGALTRIIVTTBAIRDKLVIGEYORNSIPILEKAITKDT 300  
Db 241 TRPELVYVDVLRSAINTGTDEGALTRIIVTTBAIRDKLVIGEYORNSIPILEKAITKDT 300  
QY 301 RGDYKMLVALLGEDDA 317  
301 RGDYKMLVALLGEDDA 317  
Db 301 RGDYKMLVALLGEDDA 317

RESULT 2  
AAG0874  
ID AAG0874 standard; protein; 317 AA.  
XX  
AC AAG0874;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6582.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
PD  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135124P.  
PR 24-MAY-1999; 99US-0135533P.  
PR 25-MAY-1999; 99US-0135629P.  
PR 27-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145293P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.

PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160983P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 26-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 1611; DB 3; Length 317;  
Best Local Similarity 100.0%; Pred. No. 1.2e-139;  
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATLKVSDSVAPSPDADAEQLRTAFEGWGTNEEDLIISLAHRSABQRKYIRQAHETTYGED 60  
DB 1 MATLKVSDSVAPSPDADAEQLRTAFEGWGTNEEDLIISLAHRSABQRKYIRQAHETTYGED 60  
QY 61 LKKTLDKELSDPERAILLMTLEPERDALLANEATKMTSSNOQLMEVACTRTSTOLLH 120  
DB 61 LKKTLDKELSDPERAILLMTLEPERDALLANEATKMTSSNOQLMEVACTRTSTOLLH 120  
QY 121 AROAYHARYKKSLEBDVAHHTTGDPKRLIVSLVTSRYREGSRVNMTLAKOKAKVHEKIX 180  
DB 121 AROAYHARYKKSLEBDVAHHTTGDPKRLIVSLVTSRYREGSRVNMTLAKOKAKVHEKIX 180  
QY 161 DGHVDEDEVIRILSTRSAQINATPNRYODDHGEIILKSLBEGDDDDFFLALLRSTIOCL 240  
DB 161 DGHVDEDEVIRILSTRSAQINATPNRYODDHGEIILKSLBEGDDDDFFLALLRSTIOCL 240  
QY 241 TRPELYPVDVLRSAINKGTDEGALTRIVTTRAEIDLNVIGEYGRNRSIPLEKAITROT 300  
DB 241 TRPELYPVDVLRSAINKGTDEGALTRIVTTRAEIDLNVIGEYGRNRSIPLEKAITROT 300  
QY 301 RGDYERKMVLALGEDDA 317  
DB 301 RGDYERKMVLALGEDDA 317

RESULT 3  
ID ADJ63872 standard; protein; 317 AA.  
AC ADJ63872;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Plant lipid metabolism protein OO-10 SEQ ID NO:74.  
XX  
KW plant; lipid metabolism protein; LMP; seed storage compound;  
XX  
OS Unidentified.  
XX  
OS WO2004013304-A2.  
XX  
PN  
XX  
PD 12-FEB-2004.  
XX  
PD 04-AUG-2003; 2003WO-US024364.  
XX  
PF  
XX 02-AUG-2002; 2002US-0400803P.

XX (BADI ) BASF PLANT SCI GMBH.  
 PA Miltendorf V, Haertel HA, Bauer J, Oswald O;  
 PI WPI; 2004-157121/15.  
 DR N-PSDB; ADJ63871.  
 XX New lipid metabolism proteins and nucleic acids, useful in producing  
 PT transgenic plants with increased levels of seed storage compound, e.g.  
 PT lipid, a fatty acid, a starch or a seed storage protein.  
 XX Claim 1; SEQ ID NO 74; 115bp; English.  
 PS  
 XX The invention relates to novel isolated lipid metabolism proteins (LMP)  
 CC and encoding nucleic acids comprising a polynucleotide sequence encoding  
 CC a polypeptide that functions as a modulator of seed storage compounds in  
 CC a plant. The LMP nucleic acid is useful in producing transgenic plants  
 CC with increased levels of seed storage compound, e.g. lipid, a fatty acid,  
 CC a starch or a seed storage protein, as markers for specific regions of  
 CC the genome and for evolutionary and protein structural studies. The  
 CC present sequence represents an LMP of the invention.  
 CC  
 XX Sequence 317 AA:  
 SQ  
 Query Match 100.0%; Score 1611; DB 8; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-139; Mismatches 0; Gaps 0;  
 Matches 317; Conservative 0; Indels 0; Gaps 0;  
 QY 1 MATLKVSQSVAPSDAEQRLTAPEGWGTNEDLIISILAHRSAEQRKVIROAVHETGSD 60  
 DB 1 MATLKVSQSVAPSDAEQRLTAPEGWGTNEDLIISILAHRSAEQRKVIROAVHETGSD 60  
 QY 61 LKTTLDKEISNDPERAILLWTLPEGERDALLANEATKRWTSNQVLMVACTRTSTQLH 120  
 DB 61 LKTTLDKEISNDPERAILLWTLPEGERDALLANEATKRWTSNQVLMVACTRTSTQLH 120  
 QY 121 AROAVHARVYKSLSEEDVAHHHTGDFRKLVSIVTSYRYEGDEVNMTLAKOEKLVHEKIK 180  
 DB 121 AROAVHARVYKSLSEEDVAHHHTGDFRKLVSIVTSYRYEGDEVNMTLAKOEKLVHEKIK 180  
 QY 181 DKHYNDEDEVIRIILSTRSKAQINATFNRYODHGEIILSLSEGGDDDKFALLRSTIOCL 240  
 DB 181 DKHYNDEDEVIRIILSTRSKAQINATFNRYODHGEIILSLSEGGDDDKFALLRSTIOCL 240  
 QY 241 TRPELYFVDVLRSAINKTGDEGALTRIVTTTTRABIDLKVIAGEYORRNSIPLEKAITKOT 300  
 DB 241 TRPELYFVDVLRSAINKTGDEGALTRIVTTTTRABIDLKVIAGEYORRNSIPLEKAITKOT 300  
 QY 301 RGDYERKMLVALLGEDDA 317  
 DB 301 RGDYERKMLVALLGEDDA 317

XX 20-OCT-2003; 2003WO-EP011658.  
 PF  
 XX 18-OCT-2002; 2002EP-00079408.  
 PR  
 XX (CROP-) CROPPDESIGN NV.  
 PA  
 XX Inze D, De Veylder L, Vlieghe K;  
 PI WPI; 2004-348466/32.  
 DR N-PSDB; ADN74566.  
 XX  
 XX Altering plant characteristics, useful for producing plants for enzyme or  
 PT pharmaceutical production comprises modifying in a plant, expression of  
 PT one or more nucleic acids and/or modifying level or activity of one or  
 PT more proteins.  
 XX Claim 1; SEQ ID NO 2462; 134bp; English.  
 PS  
 XX This invention relates to a novel method for altering one or more plant  
 CC characteristics. Specifically, it refers to identifying genes that are up  
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
 CC alter plant characteristics accordingly. The present invention describes  
 CC generating transgenic plants for the production of growth regulators,  
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
 CC the altered plant characteristics are selected from increased yield or  
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
 CC or physiology, altered lipid mobilization, blochchemistry, signal  
 CC transduction, storage lipid mobilization and/or altered photosynthesis,  
 CC each relative to the corresponding wild type plants. Accordingly, these  
 CC sequences can also be useful as positive or negative selectable markers  
 CC during transformation of cells or tissues. The identified genes play a  
 CC role in a variety of biological processes such as DNA replication, cell  
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
 CC transcription factors. This polypeptide sequence is thale cress protein  
 CC expressed by a gene repressed 1.3 fold or more in plants overexpressing  
 CC the E2Fa/Dpa transcription factor, given in an exemplification of the  
 CC invention.  
 CC  
 XX Sequence 317 AA:  
 SQ  
 Query Match 100.0%; Score 1611; DB 8; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-139; Mismatches 0; Gaps 0;  
 Matches 317; Conservative 0; Indels 0; Gaps 0;  
 QY 1 MATLKVSQSVAPSDAEQRLTAPEGWGTNEDLIISILAHRSAEQRKVIROAVHETGSD 60  
 DB 1 MATLKVSQSVAPSDAEQRLTAPEGWGTNEDLIISILAHRSAEQRKVIROAVHETGSD 60  
 QY 61 LKTTLDKEISNDPERAILLWTLPEGERDALLANEATKRWTSNQVLMVACTRTSTQLH 120  
 DB 61 LKTTLDKEISNDPERAILLWTLPEGERDALLANEATKRWTSNQVLMVACTRTSTQLH 120  
 QY 121 AROAVHARVYKSLSEEDVAHHHTGDFRKLVSIVTSYRYEGDEVNMTLAKOEKLVHEKIK 180  
 DB 121 AROAVHARVYKSLSEEDVAHHHTGDFRKLVSIVTSYRYEGDEVNMTLAKOEKLVHEKIK 180  
 QY 181 DKHYNDEDEVIRIILSTRSKAQINATFNRYODHGEIILSLSEGGDDDKFALLRSTIOCL 240  
 DB 181 DKHYNDEDEVIRIILSTRSKAQINATFNRYODHGEIILSLSEGGDDDKFALLRSTIOCL 240  
 QY 241 TRPELYFVDVLRSAINKTGDEGALTRIVTTTTRABIDLKVIAGEYORRNSIPLEKAITKOT 300  
 DB 241 TRPELYFVDVLRSAINKTGDEGALTRIVTTTTRABIDLKVIAGEYORRNSIPLEKAITKOT 300  
 QY 301 RGDYERKMLVALLGEDDA 317  
 DB 301 RGDYERKMLVALLGEDDA 317



ID AAG09199 standard; protein; 327 AA.  
XX AAG09199;  
AC  
XX 17-OCT-2000 (first entry)  
DT  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 7038.  
DE  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
OS Arabidopsis thaliana.  
XX EP1033405-A2.  
PN  
XX 06-SEP-2000.  
PD  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
PF  
XX  
XX 25-FEB-1999; 99US-0121625P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134321P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134941P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144633P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148131P.  
PR 12-AUG-1999; 99US-0148345P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.

```
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159283P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-016120P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 1 100.0%; Score 1611; DB 3; Length 327;  
Best Local Similarity 100.0%; Pred. No. 1,3e-139;  
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MATLKXSDSVPAQSDAEQRTAFEGWGTNEDLIISILHRSAEQRKVIROAVHEHYGGD 60
DB 11 MATLKXSDSVPAQSDAEQRTAFEGWGTNEDLIISILHRSAEQRKVIROAVHEHYGGD 70
QY 61 LKLTLDKELSNDFERAIIILWTLPEGERDALLANEATKRWTSNOQVIMEVACTRTSTQLH 120
DB 71 LKLTLDKELSNDFERAIIILWTLPEGERDALLANEATKRWTSNOQVIMEVACTRTSTQLH 130
QY 121 ARQAVYARVKKSLSEEDVAHHTTDFRKLAVSLVTSRYEDEDVNMTLAQEALVHEKIK 180
DB 131 ARQAVYARVKKSLSEEDVAHHTTDFRKLAVSLVTSRYEDEDVNMTLAQEALVHEKIK 190
QY 181 DKHYNDEVIRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALLRSTIOCL 240
```

```
DB 191 DKHYNDEVIRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALLRSTIOCL 250
QY 241 TRPELYFVUVLMSAINKTGTDBGALTRIVTTTAEIDLKITGEHYQRNSIPLEKAITKOT 300
DB 251 TRPELYFVUVLMSAINKTGTDBGALTRIVTTTAEIDLKITGEHYQRNSIPLEKAITKOT 310
QY 301 RGDYKMLVALLGEEDA 317
DB 311 RGDYKMLVALLGEEDA 327

RESULT 6
AAG08873
ID AAG08873 standard; protein; 328 AA.
AC AAG08873;
XX
XX 17-OCT-2000 (first entry)
DT
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6581.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridization assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0132865P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 20-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
```

PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140911P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 19-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149425P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0158293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161932P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 1611; DB 3; Length 328;  
Best Local Similarity 100.0%; Pred. No. 1.3e-139;  
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MATLKVSQVAPSDAQLRTAFEGMGCTNEDLIISILAHRSAEQRKVIROAHYETGSD 60

```

Db      12 MATLKVSDSVPAESDDAEQLRTAFEGGNTEDLIISILHRSAGQRVIRQAYHETGYED 71
Qy      61 LKLTLDKELSNDFERRALLMTLEPERDALLANEAATKRTSSNOVLMEVACTRTSTOLH 120
Db      72 LKLTLDKELSNDFERRALLMTLEPERDALLANEAATKRTSSNOVLMEVACTRTSTOLH 131
Qy      121 AROAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRSGDEVNMTLAKQBAKLVHEKIK 180
Db      132 AROAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRSGDEVNMTLAKQBAKLVHEKIK 191
Qy      181 DKHYNDEDVIRILSTRSKAQINATFNRYODHGEIILKSLEGGDDDKFLALLRSTIOCL 240
Db      192 DKHYNDEDVIRILSTRSKAQINATFNRYODHGEIILKSLEGGDDDKFLALLRSTIOCL 251
Qy      241 TRPELVFVDVLRSAINKTGDGALTRIVTTTAEIDLVKVIAGEYQRNSIPILEKAITKDT 300
Db      252 TRPELVFVDVLRSAINKTGDGALTRIVTTTAEIDLVKVIAGEYQRNSIPILEKAITKDT 311
Qy      301 RGDYERKMLVALLGEDDA 317
Db      312 RGDYERKMLVALLGEDDA 328

RESULT 7
ID AAB16294 standard; protein; 316 AA.
AC AAB16294;
DT 31-OCT-2000 (first entry)
XX
DE Eucalyptus grandis amexin protein sequence SEQ ID NO:56.
XX
KW Eucalyptus grandis; pinus radiata; Monterey pine; modification;
KM plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
XX transgenic plant.
OS Eucalyptus grandis.
XX
PN MO200022092-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99MO-NZ000169.
XX
PR 13-OCT-1998; 98US-00170862.
XX
PR 11-AUG-1999; 99US-0148426P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Bloksberg LN;
XX
DR WPI; 2000-339328/29.
XX N-PSDB; AAA67100.
XX
PT New genes encoding proteins involved in a plant polysaccharide
XX biosynthetic pathway, useful for modulating or altering the
XX polysaccharide content, composition or structure of the plant.
XX
PS Claim 17, Page 66, 301pp, English.
XX
CC The present invention describes isolated polynucleotides (PN) comprising
XX a sequence selected from one of 835 nucleotide sequences given in
XX AAA67073 to AAA67907, their (reverse) complements, sequences producing an
XX Expectation (E) value of 0.01 or less compared to the 835 sequences
XX CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
XX CC 835 sequences or sequences that are degenerately equivalent or allelic to
XX CC the 835 sequences. The polynucleotides are used to modify the activity of
XX CC a polypeptide involved in a polysaccharide biosynthetic pathway in the
XX CC plant. They are especially used to modulate or alter the polysaccharide
XX CC content, composition or structure of the plant. AAB16266 to AAB16340 are

```

```

CC protein encoded by some of the polynucleotide sequence given in the
XX present invention
XX
SQ Sequence 316 AA;
Query Match 73.6%; Score 1185.5; DB 3; Match 316;
Best Local Similarity 71.9%; Pred. No. 1.9e-100;
Matches 228; Conservative 38; Mismatches 50; Indels 1; Gaps 1;

Qy      1 MATLKVSDSVPAESDDAEQLRTAFEGGNTEDLIISILHRSAGQRVIRQAYHETGYED 60
Db      1 MSTLTVPQPLPVADDCQELRTAFAGMGTEKLIISILGHRNAQQRKLIRQYAEFTGYED 60
Qy      61 LKLTLDKELSNDFERRALLMTLEPERDALLANEAATKRTSSNOVLMEVACTRTSTOLH 120
Db      61 LKLTLDKELSNDFERRALLMTLEPERDALLANEAATKRTSSNOVLMEVACTRTSTOLH 120
Qy      121 AROAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRSGDEVNMTLAKQBAKLVHEKIK 180
Db      121 AROAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRSGDEVNMTLAKQBAKLVHEKIK 180
Qy      181 DKHYNDEDVIRILSTRSKAQINATFNRYODHGEIILKSLEGGDDDKFLALLRSTIOCL 240
Db      181 EKAYGHEDLIRILATRSKAQVNAATLNHYKNERGNDINKDKT-DPKDAFLTLIRATVKCL 239
Qy      241 TRPELVFVDVLRSAINKTGDGALTRIVTTTAEIDLVKVIAGEYQRNSIPILEKAITKDT 300
Db      240 TRPEKFEKVLRLAINKRTGDGALTRIVVATRAEVDMKFISEYQRNSIPIEDRAIVKDT 299
Qy      301 RGDYERKMLVALLGEDDA 317
Db      300 TGDYERKMLALIGHVEA 316

RESULT 8
ID AAB16321 standard; protein; 316 AA.
AC AAB16321;
DT 31-OCT-2000 (first entry)
XX
DE pinus radiata amexin protein sequence SEQ ID NO:114.
XX
KW Eucalyptus grandis; pinus radiata; Monterey pine; modification;
KM plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
XX transgenic plant.
OS pinus radiata.
XX
PN MO200022092-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99MO-NZ000169.
XX
PR 13-OCT-1998; 98US-00170862.
XX
PR 11-AUG-1999; 99US-0148426P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Bloksberg LN;
XX
DR WPI; 2000-339328/29.
XX N-PSDB; AAA67128.
XX
PT New genes encoding proteins involved in a plant polysaccharide
XX biosynthetic pathway, useful for modulating or altering the
XX polysaccharide content, composition or structure of the plant.
XX
PS Claim 17, Page 94-95, 301pp, English.
XX
CC

```



PF 14-AUG-2002; 2002US-00219220.  
 XX  
 PR 04-JUN-1999; 99US-00325932.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 XX  
 PI Film B, Lasham A;  
 XX  
 DR WPI; 2003-786916/74.  
 XX  
 PT New isolated polynucleotide useful for modulating programmed cell death,  
 PT altering the development cycle of plant cells, and subsequently modifying  
 PT plant development.  
 XX  
 PS Claim 12; Page 47-48; 21app; English.  
 XX  
 CC The invention describes an isolated polynucleotide (1) comprising a  
 CC nucleotide sequence that is one of 145 fully defined sequences of 221-  
 CC 3415 base pairs (bp), given in the specification, and/or its complements,  
 CC reverse complements, reverse sequences, or sequences having 75, 90 or 95  
 CC % sequence identity to or that hybridise under stringent hybridisation  
 CC conditions to one of the 145 sequences. The methods and compositions of  
 CC the present invention to do with (1) are useful for modulating programmed  
 CC cell death and thereby altering the development cycle of plant cells, and  
 CC altering plant development. This is the amino acid sequence of a protein  
 CC associated with the programmed cell death pathway.  
 XX  
 SQ Sequence 316 AA;  
 Query Match 72.7%; Score 1171.5; DB 7; Length 316;  
 Best Local Similarity 71.3%; Pred. No. 3.8e-99;  
 Matches 226; Conservative 39; Mismatches 51; Indels 1; Gaps 1;  
 QY 1 MATLKVS DSPAPSDDAEQLRTAFEGWGTNEDLIISILAHRSAEQRKVIRQAYHETGYED 60  
 DB 1 MSTLTVPQPLPPVADDEQLRTAFAGWGTEKLIISILHRRNAQRKLIQTYAETGYED 60  
 QY 61 LKTLIDKELSDNPEERAILMTLEPGERDALLANEATKRTSSNOVLMETVACTRTSTQOLH 120  
 DB 61 LKRALDKELSDNPEERAILMTLEPGERDALLANEATKRTSSNOVLMETVACTRTSTQOLH 120  
 QY 121 AROAYHARYKKSLEEDVAHHTTGDFRKLIVSLVTSYRGEDEVNMTLAKQBAKLVHEKIK 180  
 DB 121 AROAYHARYKKSLEEDVAHHTTGDFRKLIVPLGSSYRNDDEVNMTLAKQBAKILHEKIS 180  
 QY 181 DKHYNDEDEVIRILSTRSKAQINATFNRYQDHGEELIKSLSEGGDDDKFLALLRSTIQCL 240  
 DB 181 EKAYGHEDLIRILSTRSKAQVNATFNRYKNEFGNDINKDLKT-DPKDAFLTIIRATVYKCL 239  
 QY 241 TREPELVFVDVLRASINKTGDEGALTRIVTTTAEIDKVIAGEYQRNSIPLEKAITKOT 300  
 DB 240 TREPEKTFEKVLRALINRGTDGALTRVAVTRADEVDMKFTSEYQRNSIPLEKAIYKOT 299  
 QY 301 RGDYKMLVALLEGDDA 317  
 DB 300 TGDYKMLLALIGHVEA 316  
 Db  
 RESULT 11  
 ID ADB94896  
 AD B94896 standard; protein; 316 AA.  
 XX ADB94896;  
 XX  
 DE 04-DEC-2003 (first entry)  
 XX  
 XX Programmed cell death pathway protein annexin #6.  
 XX  
 KM programmed cell death; plant development; plant cell cycle; ATL2; DAD1;  
 KM Dnase; 11s; 1sdi; nucellin-like aspartic protease; annexin; prohibitin;  
 KM fan-like protein; rac2; retinoblastoma-related protein; SIVA; TFIID;  
 KM TEGT; xylogenic RNase; pur-alpha; cyteine protease; Rps5-like protein;  
 KM GP 91 NADPH oxidase subunit; NPR-like protein; BAG-1;

KM defender against cell death; lethal leaf spot; lesion stimulating death;  
 KM seven in absentia; transcription initiation factor;  
 KM testis enhanced gene transcript.  
 XX  
 OS Eucalyptus grandis.  
 XX  
 PN US2003082724-A1.  
 XX  
 PD 01-MAY-2003.  
 XX  
 PF 14-AUG-2002; 2002US-00219220.  
 XX  
 PR 04-JUN-1999; 99US-00325932.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 XX  
 PI Film B, Lasham A;  
 XX  
 DR WPI; 2003-786916/74.  
 DR N-PSDB; ADB94854.  
 XX  
 PT New isolated polynucleotide useful for modulating programmed cell death,  
 PT altering the development cycle of plant cells, and subsequently modifying  
 PT plant development.  
 XX  
 PS Claim 12; Fig 27; 21app; English.  
 XX  
 CC The invention describes an isolated polynucleotide (1) comprising a  
 CC nucleotide sequence that is one of 145 fully defined sequences of 221-  
 CC 3415 base pairs (bp), given in the specification, and/or its complements,  
 CC reverse complements, reverse sequences, or sequences having 75, 90 or 95  
 CC % sequence identity to or that hybridise under stringent hybridisation  
 CC conditions to one of the 145 sequences. The methods and compositions of  
 CC the present invention to do with (1) are useful for modulating programmed  
 CC cell death and thereby altering the development cycle of plant cells, and  
 CC altering plant development. This is the amino acid sequence of a protein  
 CC associated with the programmed cell death pathway.  
 XX  
 SQ Sequence 316 AA;  
 Query Match 70.3%; Score 1132.5; DB 7; Length 316;  
 Best Local Similarity 68.8%; Pred. No. 1.5e-95;  
 Matches 218; Conservative 48; Mismatches 50; Indels 1; Gaps 1;  
 QY 1 MATLKVS DSPAPSDDAEQLRTAFEGWGTNEDLIISILAHRSAEQRKVIRQAYHETGYED 60  
 DB 1 MATIAVPPSPVSPAPSDAEQLRTAFAGWGTEKLIISILHRRNAQRKLIQTYAETGYED 60  
 QY 61 LKTLIDKELSDNPEERAILMTLEPGERDALLANEATKRTSSNOVLMETVACTRTSTQOLH 120  
 DB 61 LKRALDKELSDNPEERAILMTLEPGERDALLANEATKRTSSNOVLMETVACTRTSSMELPM 120  
 QY 121 AROAYHARYKKSLEEDVAHHTTGDFRKLIVSLVTSYRGEDEVNMTLAKQBAKLVHEKIK 180  
 DB 121 VROAYHARYKKSLEEDIAHHTTGDFRKLIVPLASTFRYGPVAVNMTLARSSEKILHEKIH 180  
 QY 181 DKHYNDEDEVIRILSTRSKAQINATFNRYQDHGEELIKSLSEGGDDDKFLALLRSTIQCL 240  
 DB 181 EKAYNHDELIRIVTTRSKAQVNATFNRYKNEFGNAINKDL-KADPNDEFLKILRSKIKCL 239  
 QY 241 TREPELVFVDVLRASINKTGDEGALTRIVTTTAEIDKVIAGEYQRNSIPLEKAITKOT 300  
 DB 240 TYPEKTFEKVLRALINKTGDEGALTRVAVTRADEVDMKFTSEYQRNSIPLEKAIYKOT 299  
 QY 301 RGDYKMLVALLEGDDA 317  
 DB 300 SGDYKMLLALVGHEDA 316  
 Db  
 RESULT 12  
 ID AAG09201  
 AA G09201 standard; protein; 211 AA.  
 XX

AC AAG09201;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 7040.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132837P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 18-MAY-1999; 99US-0134370P.  
PR 19-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137502P.  
PR 04-JUN-1999; 99US-0137528P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144881P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.



```
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152353P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 66.5%; Score 1071; DB 3; Length 211;  
Best Local Similarity 100.0%; Pred. No. 3.8e-90; Indels 0; Gaps 0;  
Matches 211; Conservative 0; Mismatches 0;

```
QY 107 MBVACTRTSTOLHARQAYHARYKKSLEEDVAHHTTGDFPKKLVSIVTSYRYGDEVNMT 166
DB 1 MBVACTRTSTQLHARQAYHARYKKSLEEDVAHHTTGDFPKKLVSIVTSYRYGDEVNMT 60
QY 167 LAQOAKLVHEKIKDKKYNDEVDVIRIISTSKAQINATFNRYODDHGEETILKSLERGGDD 226
DB 61 LAQOAKLVHEKIKDKKYNDEVDVIRIISTSKAQINATFNRYODDHGEETILKSLERGGDD 120
QY 227 DKFLALRSTIOCTTRELTFVNDLRSAINKTGDSGALFRIYTTTAEIDLKVIAGEYOR 286
DB 121 DKFLALRSTIOCTTRELTFVNDLRSAINKTGDSGALFRIYTTTAEIDLKVIAGEYOR 180
QY 287 RNSIPLEKAITKOTRGDYEKMLVALIGEDDA 317
DB 181 RNSIPLEKAITKOTRGDYEKMLVALIGEDDA 211
```

```
RESULT 13
AAG08875
ID AAG08875 standard; protein; 211 AA.
XX
AC AAG08875;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6583.
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-01231825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-013149P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135533P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
```

PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	99US-0139839P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140685P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144335P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144684P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145092P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145192P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161359P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146389P.	PR	26-OCT-1999;	99US-0161361P.
PR	03-AUG-1999;	99US-0147038P.			

Qy 287 NSIPIEKAITKDTGRGPEKKLVALLGDDA 317  
Db 181 NSIPIEKAITKDTGRGPEKKLVALLGDDA 211

RESULT 14  
AAC06311  
ID AAC06311 standard; protein: 211 AA.

AC AAC06311;  
XX 17-OCT-2000 (first entry)  
XX  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 3040.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0030139.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.

PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-014187P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142290P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145813P.  
PR 27-JUL-1999; 99US-0145818P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.

PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151033P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161202P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 66.5%; Score 1071; DB 3; Length 211;  
Best Local Similarity 100.0%; Fred.No.3.8e-90;  
Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 MEVACTRTSTQTLHAQAQVARYKKSLEEDVAHHTTGDPFRKLLVSLVTSYRYEGDEVNMT 166  
DB 1 MEVACTRTSTQTLHAQAQVARYKKSLEEDVAHHTTGDPFRKLLVSLVTSYRYEGDEVNMT 60  
QY 167 LAKQEAKLVEHKIKDKHYNDEVDVIRILSTRSKAQINATFNRYODDHGEIILKSLLEGDD 226

DB 61 LAKQEAKLVEHKIKDKHYNDEVDVIRILSTRSKAQINATFNRYODDHGEIILKSLLEGDD 120  
QY 227 DKFLALRSTIOCLRPPELXFVDVVRSAINKGTDEGALTRIVTTRARIDLVIGEYOR 286  
DB 121 DKFLALRSTIOCLRPPELXFVDVVRSAINKGTDEGALTRIVTTRARIDLVIGEYOR 180  
QY 287 RNSIPLEKAITKDTRGDYEKMLVALLGEDDA 317  
DB 181 RNSIPLEKAITKDTRGDYEKMLVALLGEDDA 211  
RESULT 15  
AAG23493  
ID AAG23493 standard; protein; 317 AA.  
XX  
AC AAG23493;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 26821.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137223P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 22-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144335P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147202P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148371P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149375P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0152363P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157153P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 66.2%; Score 1067; DB 3; Length 317;  
Best local Similarity 64.0%; Pred. No. 1.6e-89;  
Matches 203; Conservative 50; Mismatches 64; Indels 0; Gaps 0;  
QY 1 MATLKVSQSDVPAPSDABQRTAFQGWGCTNEDLIISILAHRAEQKRYIRQAYHETGSD 60

Db 1 MASLKVPSPVPLPEDDAEQJHKAFGSGWGTNEKLIISILAHRNAAQBSLIRSVYAATYNED 60  
QY 61 LKTIIDKETSNDPERAILMTLEPGERDALLANEATKRWTSNOVIMEVACTRTSTQLH 120  
Db 61 LKALDKELSSDERAVMLTLDPPERDAYLAKESTKMFYKNNWLVVEIACRPALELIK 120  
QY 121 ARQAYHARYKKSLEEDVAHHTTGDFFRKLIVSLVTSYREGEDEVNMTLAKOEAKLVHEKIK 180  
Db 121 VKQAYQARYKKSIEEDVAQHTSGDLKLLPLVSTFRYEGDDVNMMLARSEAKILHEKVS 180  
QY 181 DKRYNDEVIIRIISTRSKAQINATFNRYODHGEEILKSLEEGDDDKFLALLRSTIQCL 240  
Db 181 EKSYSDDDFRILYTRSKAQIGATLNHNNNEYGNAINKNLKEESDDNDYMKLLRAVITCL 240  
QY 241 TREPELVFVULRSAINKTGTDEGALTRIVTTRAEIDLKVIGEYQRRNSIPLEKATKDT 300  
Db 241 TYPEKHFEKVLRLSINKMGTDWGLTRVVTTRTEVDMERIKEEYQRRNSIPLDRAIYKDT 300  
QY 301 RGDYERKLVALLGEDDA 317  
Db 301 SGDYEDMLVALLGHGDA 317

Search completed: August 19, 2005, 20:40:41  
Job time : 169 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 20:36:40 ; Search time 40 Seconds  
(without alignments)  
762,518 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611  
Sequence: 1 MATLKVS DSVAPASDAEQL.....KOTRGDYKKLVALLGEDDA 317

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0 %  
Maximum Match 100 %  
Listing first 45 summaries

Database :

1: PIR 79: \*  
2: PIR1: \*  
3: PIR2: \*  
4: PIR3: \*  
5: PIR4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1611	100.0	317	2 C86479	probable annexin p
2	1202.5	74.6	316	2 T31428	fiber annexin - up
3	1128.5	70.0	315	2 T10805	annexin - upland c
4	1117.5	69.4	314	2 S66274	annexin - pepper
5	1107	68.7	315	2 T10807	annexin 2 - upland
6	1080	67.0	315	2 T06322	annexin, isoform P
7	1067.5	66.3	308	2 T09552	annexin - alfalfa
8	1037.5	64.4	316	2 T50027	annexin-like prote
9	1000.5	62.1	314	2 T02961	annexin p33 - maiz
10	991.5	61.5	318	2 T50026	annexin-like prote
11	949.5	58.9	314	2 T02975	annexin p35 - maiz
12	772.5	48.0	271	2 S56674	annexin homolog R
13	702.5	37.4	321	2 A84809	probable annexin I
14	523	32.5	316	2 B96704	probable annexin T
15	501	31.1	316	1 LUHUS	annexin XIII, inte
16	496	30.8	316	2 A57076	annexin XIIIa - do
17	496	30.8	357	2 B57076	annexin XIIIb - do
18	472	29.3	323	1 LUHUS	annexin III - huma
19	462	28.7	319	2 H84808	probable annexin I
20	461.5	28.6	673	1 AOHU68	annexin VI [valida
21	454	28.2	324	1 LUR73	annexin III - rat
22	451	28.0	505	2 A5152	annexin XI - huma
23	450	27.9	503	1 LUBO11	annexin XI form A
24	450	27.9	503	1 S23447	annexin XI form B
25	448.5	27.7	673	1 S01786	annexin VI - mouse
26	447	27.7	488	1 LUHU7	annexin VII, long
27	444	27.6	503	1 LURB11	annexin XI - rabbi
28	435.5	27.0	318	1 LURP4	annexin IV - pig
29	434	26.9	463	2 S29170	annexin VII - mous

30	429.5	26.7	673	1 S52844	annexin VI - rat
31	425.5	26.4	321	1 A42077	annexin IV - human
32	424.5	26.4	671	2 UC2029	annexin - chicken
33	421.5	26.2	319	1 LUBO4	annexin IV - bovin
34	413.5	25.7	321	1 LUFF10	annexin X - fruit
35	408.5	25.4	512	2 S70644	annexin VII - Afri
36	407.5	25.3	321	1 LUCH5	annexin V - chicke
37	405	25.1	340	1 JQ1297	annexin II type 1
38	403.5	25.0	339	1 LUBO36	annexin II - bovin
39	402	25.0	339	1 LUCH2	annexin II - chick
40	398.5	24.7	339	1 LUBH36	annexin II - human
41	398.5	24.7	340	1 JQ1298	annexin II type 2
42	396.5	24.6	339	1 LUMS36	annexin II - mouse
43	396	24.6	346	1 LUR71	annexin I - rat
44	393	24.4	346	1 LURU	annexin I - human
45	392.5	24.4	327	1 LUBH8	annexin VII - hum

ALIGNMENTS

RESULT 1

C86479

probable annexin protein - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear creess)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C/Accession: C86479

R/Theologian, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

A.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marziani

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shum, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; WUID:21016719; PMID:11130712

A/Accession: C86479

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-317 <STO>

A/Cross-references: UNIPROT:Q9SYT0; GB:AE005172; NID:98778967; PIDN:AAF79882.1; GSPDB:GI

C/Genetics:

A/Map position: 1

C/Superfamily: annexin I; annexin repeat homology

Query Match 100.0%; Score 1611; DB 2; Length 317;

Best Local Similarity 100.0%; Pred. No. 1.6e-97; Mismatches 0; Indels 0; Gaps 0;

Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MATLKVS DSVAPASDAEQLRTAFEGWGTNEDLISILHRSAPORKVIRQAYHETGYED 60		
DB	1	MATLKVS DSVAPASDAEQLRTAFEGWGTNEDLISILHRSAPORKVIRQAYHETGYED 60		
QY	61	LKLTDLKELSNDFERRAILMTLEPGERDALLANEATKRWTSNQVIMEVAICTSTYOLH 120		
DB	61	LKLTDLKELSNDFERRAILMTLEPGERDALLANEATKRWTSNQVIMEVAICTSTYOLH 120		
QY	121	ARQAYHARKYKSLSEBVVAHHTTGDPRKLVSVTSYRYSDEVDNMTLAKOEAALVHEKIK 180		
DB	121	ARQAYHARKYKSLSEBVVAHHTTGDPRKLVSVTSYRYSDEVDNMTLAKOEAALVHEKIK 180		
QY	181	DKHYNDEVDVIRILSTRSKAQINATFNRYODDHEEILKSLSEGGDDDKFLALRSTIOCL 240		
DB	181	DKHYNDEVDVIRILSTRSKAQINATFNRYODDHEEILKSLSEGGDDDKFLALRSTIOCL 240		
QY	241	TRPELYFVDVLRSAINKTGTDEGALTRIVTTAEIDLKVIIGEYQRRNSIPLKAIKTOT 300		
DB	241	TRPELYFVDVLRSAINKTGTDEGALTRIVTTAEIDLKVIIGEYQRRNSIPLKAIKTOT 300		
QY	301	RGDYKKLVALLGEDDA 317		



Db 301 RGDYKMLVALLGEDDA 317

RESULT 2

TJ31428

fiber annexin - upland cotton

C/Species: Gossypium hirsutum (upland cotton)

C/Date: 29-Oct-1999 #sequence\_rev15ion 29-Oct-1999 #text\_change 09-Jul-2004

C/Accession: TJ31428

R/Shin, H.; Brown, R.M.

submitted to the EMBL Data Library, February 1997

A/Reference number: Z21028

A/Accession: TJ31428

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-316 <SHI>

A/Cross-references: UNIPROT:P93157; EMBL:U99609; NID:g3493171; PID:g3493172

A/Experimental source: Strain Texas Marker1

C/Superfamily: annexin I; annexin repeat homology

C/Keywords: calcium binding

Query Match 74.6%; Score 1202.5; DB 2; Length 316;

Best Local Similarity 73.5%; Pred. No. 5.3e-71;

Matches 230; Conservative 37; Mismatches 45; Indels 1; Gaps 1;

Db 1 MATLKVSQSVAPSPDADAEQRTAFEGMGNTEDLIISILAHSAEQRKVIROAYHETYGED 60

1 MATLVPTTVPSVSEDCQQRKAFSGMGNTGIIIDILGRNAEQRLIRKTYAETYGED 60

QY 61 LKTLDEKLSNDFERAILMTLEPERDALLANEATKRWTSNOVLMVACTRTSTQLLH 120

Db 61 LKALDKELSNDFERAILMTLEPERDALLANEATKRWTSNOVLMVACTRTSTQLLH 120

QY 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAQVHEKIK 180

Db 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAQVHEKIK 180

QY 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAQVHEKIK 180

Db 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAQVHEKIK 180

QY 181 DKRYNDEVYIRILSTRSKAQINATFNRYODDHEGEEILKSLBEGDDDDKFLALRSTIOCL 240

Db 181 DKRYNDEVYIRILSTRSKAQINATFNRYODDHEGEEILKSLBEGDDDDKFLALRSTIOCL 240

QY 241 TRPELVFVDVLRSAINKGTDEGALTRIVTTTTRAEIDLVIGEYQRRNSIPLEKAITKOT 300

Db 240 VYPEKFEKVLRLAIRRGTDGALTRVCTRAEVDLKIADYEQRRNSVPLTRALVKDT 299

QY 301 RGDYKMLVALLG 313

Db 300 RGDYKMLVALLG 312

RESULT 3

T10805

annexin - upland cotton (fragment)

C/Species: Gossypium hirsutum (upland cotton)

C/Date: 16-Jul-1999 #sequence\_rev15ion 16-Jul-1999 #text\_change 09-Jul-2004

C/Accession: T10805

R/Potikha, T. S.; Delmer, D. P.

submitted to the EMBL Data Library, October 1996

A/Reference number: Z17161

A/Accession: T10805

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-315 <POT>

A/Cross-references: UNIPROT:P93157; EMBL:U73746; NID:g1843524; PID:g1843525

A/Experimental source: strain Acala SJ-2; cotton fiber

C/Genetics:

A/Note: Aungh1

C/Superfamily: annexin I; annexin repeat homology

C/Keywords: calcium binding

F:85-156/Domain: annexin repeat homology <AXR>

Query Match 70.0%; Score 1128.5; DB 2; Length 315;

Best Local Similarity 69.2%; Pred. No. 3.4e-66;

Matches 216; Conservative 40; Mismatches 55; Indels 1; Gaps 1;

QY 2 ATLKVSDSVAPSPDADAEQRTAFEGMGNTEDLIISILAHSAEQRKVIROAYHETYGED 61

Db 1 ATLVTPTTVPSVSEDCQQRKAFSGMGNTGIIIDILGRNAEQRLIRKTYAETYGED 60

QY 62 LKTLDEKLSNDFERAILMTLEPERDALLANEATKRWTSNOVLMVACTRTSTQLLH 121

Db 61 LKALDKELSNDFERAILMTLEPERDALLANEATKRWTSNOVLMVACTRTSTQLLH 120

QY 122 ROAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAQVHEKIK 181

Db 121 ROAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAQVHEKIK 180

QY 182 KRYNDEVYIRILSTRSKAQINATFNRYODDHEGEEILKSLBEGDDDDKFLALRSTIOCL 241

Db 181 KRYNDEVYIRILSTRSKAQINATFNRYODDHEGEEILKSLBEGDDDDKFLALRSTIOCL 240

QY 242 TRPELVFVDVLRSAINKGTDEGALTRIVTTTTRAEIDLVIGEYQRRNSIPLEKAITKOT 301

Db 240 VYPEKFEKVLRLAIRRGTDGALTRVCTRAEVDLKIADYEQRRNSVPLTRALVKDT 299

QY 302 RGDYKMLVALLG 313

Db 300 RGDYKMLVALLG 311

RESULT 4

S66274

annexin - pepper

C/Species: Capsicum annuum (pepper)

C/Date: 15-Feb-1997 #sequence\_rev15ion 13-Mar-1997 #text\_change 09-Jul-2004

C/Accession: S66274

R/Proust, J.; Houline, G.; Schantz, M.L.; Schantz, R.

FEBS Lett. 383, 208-212, 1996

A/Title: Characterization and gene expression of an annexin during fruit development in

A/Reference number: S66274; MUID:96198603; PMID:8925897

A/Accession: S66274

A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-314 <PRO>

A/Cross-references: UNIPROT:Q42657; EMBL:X93308; NID:g1071659; PIRN:CA63710.1; PID:g107

A/Note: The sequence of residues 155-230 is shown in line for annexin (alfalfa) rather t

C/Superfamily: annexin I; annexin repeat homology

F:14-85/Domain: annexin repeat homology <AX1>

F:86-157/Domain: annexin repeat homology <AX2>

F:169-238/Domain: annexin repeat homology <AX3>

F:242-313/Domain: annexin repeat homology <AX4>

Query Match 69.4%; Score 1117.5; DB 2; Length 314;

Best Local Similarity 67.1%; Pred. No. 1.7e-65;

Matches 212; Conservative 58; Mismatches 43; Indels 3; Gaps 1;

QY 1 MATLKVSQSVAPSPDADAEQRTAFEGMGNTEDLIISILAHSAEQRKVIROAYHETYGED 60

Db 1 MATLVPTTVPSVSEDCQQRKAFSGMGNTGIIIDILGRNAEQRLIRKTYAETYGED 60

QY 61 LKTLDEKLSNDFERAILMTLEPERDALLANEATKRWTSNOVLMVACTRTSTQLLH 120

Db 61 LKALDKELSNDFERAILMTLEPERDALLANEATKRWTSNOVLMVACTRTSTQLLH 120

QY 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAQVHEKIK 180

Db 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAQVHEKIK 180

QY 181 DKRYNDEVYIRILSTRSKAQINATFNRYODDHEGEEILKSLBEGDDDDKFLALRSTIOCL 240

Db 181 DKRYNDEVYIRILSTRSKAQINATFNRYODDHEGEEILKSLBEGDDDDKFLALRSTIOCL 237

QY 241 TRPELVFVDVLRSAINKGTDEGALTRIVTTTTRAEIDLVIGEYQRRNSIPLEKAITKOT 300

Db 240 VYPEKFEKVLRLAIRRGTDGALTRVCTRAEVDLKIADYEQRRNSVPLTRALVKDT 297

```
QY      301 RGDYEKMLVALLGEDD 316
          ||||| : ||||| ::
Db      298 RGDYESMLLALLGQEE 313
```

## RESULT 5

annexin 2 - upland cotton (fragment)  
 C/Species: *Gossypium hirsutum* (upland cotton)  
 C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
 C/Accession: T10807  
 R/Potikha, T. S.; Delmer, D. P.  
 submitted to the EMBL Data Library, October 1996  
 A/Reference number: Z17161  
 A/Accession: T10807  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-315 <POT>  
 A/Cross-references: UNIPROT:P93158; EMBL:U73747; NID:g1843526; P1D:g1843527  
 A/Experimental source: strain Acala S<sub>2</sub>-2; cotton fiber  
 C/Genetics:  
 A/Note: AnnGh2  
 C/Superfamily: annexin I; annexin repeat homology  
 C/Keywords: calcium binding  
 C/85-156/Domains: annexin repeat homology <AXR>

Query Match	68.7%	Score 1107	DB 2	Length 315
Best Local Similarity	69.0%	Pred. No. 8.3e-65		
Matches 218	Conservative 43	Mismatches 53	Indels 2	Gaps 2

Qy	3	TLKVSQSVAPSSDDA-QURTAFFEGGNTNEDJIIISILHRSMEQKRVQRQAHHTYEDL	61
Qy	1	TLKVPVHVSPPSDEDAEMQURKAFEGGNTNEDJIIIDILHRRNAQONSIRKYYGAYEDL	60
Db			
Qy	62	LKTLUDELSNDPERALLMTLTPERDALLNEAKTKMTSSNOVIMEVACTRISTOLHA	122
Db	61	LKLEKELTSDPERAVLLFTLPDPAEDAHLANEARKKETSNTWIMETACRSRSHELLNV	120
Qy	122	RQAVHARYKKSLEEDVAHHTTGDPRKLVLSTVSTRYSGDEVNMTLAKQZKLVHEKID	181
Qy	121	KKAVHARYKKSLEEDVAHHTTGBYRKLVLPAVSARYGGEVNMTLAKSEKILHNDISD	180
Db			
Qy	182	KATNEDVIRILSTRNSKAOINATFNRYODHGEBSLKSLBEGDDDKFLALLRSTIOCLT	241
Db	181	KATNDEEVIRIVSTRSKAOLNATLHMYNSSFNAINKOL-KADSPDEFLKTLRIVIKCLT	239
Qy	242	RPELYEVVDVLRSAINTGTDEGALTRIVYTTTRAIDIDLKIYGEYORRNSIPLEKATKOTR	301
Db	240	TPEQYIEBKLRQAINLQSGDEWALTRVYTTTRADEVMDRIKAYORRNSIPLEQALADOTS	299
Qy	302	GDYEKMLVALLGEDDA	317
Db	300	GDYEKELLALIGGDA	315

## RESULT 6

annexin, isoform p35 - tomato  
 C/Species: Lycopersicon esculentum (tomato)  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
 C/Accession: T06322; S11463; S11462  
 R/Lim, E.K.; Roberts, M.R.; Bowles, D.J.  
 J. Biol. Chem. 273, 34920-34925, 1998  
 A/Title: Biochemical characterization of tomato annexin p35: independence of calcium binding  
 A/Reference number: Z15591; MUID:99074266; PMID:9857021  
 A/Accession: T06322  
 A/Status: translated from GB/EMBL/DDBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-315 <Lim>  
 A/Cross-references: UNIPROT:O81535; EMBL:AF079231; NID:g3378203; PIDN:AAC97493.1; PID:g3378203  
 A/Experimental source: tissue type root  
 R/Smallwood, M.; Keen, J.N.; Bowles, D.J.  
 Biochem. J. 270, 157-161, 1990

**A/Title:** Purification and partial sequence analysis of plant annexins  
**A/Reference number:** S11461; MUID:90372908; PMID:2396974

A:Accession: S11463  
A:Molecule type: protein  
A:Residues: 108, 'X', 110, 'X', 112-119, 'X', 121-127, 'R', 129, 'X', 131-133, 'HI', 136-141, 'S', 142  
A:Accession: S11462  
A:Molecule type: protein  
A:Residues: 'EXN1', 175-177, 'EX', 180-181, 'X', 183-184, 'N', 186-192, 'L', 194-195, 'X', 197, 'X'  
C:Genetics:  
A:Gene: AN35  
C:Superfamily: annexin I; annexin repeat homology  
C:Keywords: calcium binding; phospholipid binding  
/86-157/Domain: annexin repeat homology <AX>

Query Match	67.0%	Score 1080;	DB 2;	Length 315;
Best Local Similarity	65.9%	Pred. No. 4.7e-63;		
Matches 209; Conservative	47;	Mismatches 55;	Indels 2;	Gaps 2

[illegible]

## RESULT 7

annexin - alfalfa (fragment)  
 C:Species: Medicago sativa (alfalfa)  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
 C:Accession: T09552  
 R:Pitrick, M.; Hirt, H.; Heberle-Bors, E.  
 A:Title: The cDNA sequence encoding an annexin from Medicago sativa.  
 A:Reference number: 216735; MUID:94286751; PMID:8016273  
 A:Accession: T09552  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-308 <PIR>  
 A:Cross-references: UNIPROT:Q42922; EMBL:X74947; NID:g512399; PIDD:CAA52903.1; PID:g512399  
 C:Genetic8:  
 A:Gene: ann  
 C:Superfamily: annexin I; annexin repeat homology  
 C:Keywords: calcium binding; phospholipid binding  
 C:/7-78/Domain: annexin repeat homology <AXR>

Query Match	66.3%	Score 1667.5;	DB 2;	Length 308;
Best Local Similarity	67.1%	Pred. No. 3e-62;		
Matches 206; Conservative	50;	Mismatches 50;	Indels 1;	Gaps 1

[illegible]

Db 63 SSDFEKAIVLMTLDPAREDAFLANQTKMLTNNISIVEIASRSPLEILKAKQAVQVR 122  
QY 130 KSLSEEDVAHHTGDPKRLIVSVTSRYEGDEVNMTLAKQEAKVHEKIKDGHYDEY 189  
Db 123 KSLSEEDVAHHTGDIRKLLVPLVGHRYEGDEVNMTLASEAKLHEKIKADAAVNHDDI 182  
QY 190 IRLSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALLRSTIOCLTRPELYFVD 249  
Db 183 IRLVTRSKAQINATFNHYNNGENVYDKDLFT-DSDEYLKILRAIKULTYPEKTFEE 241  
QY 250 VLRSAINKGTDEGALTRIVTTTAEIDLKVIAGEYGRNSIPLEKAITKOTRDYERKLV 309  
Db 242 LRLTAINKGTDENALTRVVTTRAEVDLQRIAEYGRNSVPLDRAIDKOTSGDYQKILL 301  
QY 310 ALLGEDD 316  
Db 302 ALMGHDE 308

## RESULT 8

T50027  
annexin-like protein - Arabidopsis thaliana  
N/Alternate names: protein T31P16.220  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
C/Accession: T50027  
R/Bevan, M.; Zimmermann, W.; Grueniseen, A.; Wambutt, R.; Kalicki, J.; Wohlmann, P.; St  
submitted to the Protein Sequence Database, May 2000  
A/Reference number: Z25027  
A/Accession: T50027  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-316 <BEV>  
A/Cross-references: UNIPROT:Q9LX07; EMBL:AL356332; GSPDB:GN00063; ATSP:T31P16.220  
A/Experimental source: cultivar Columbia; BAC clone T31P16  
C/Genetic:  
A/Gene: ATSP:T31P16.220  
A/Map position: 5  
A/Intons: 26/1; 74/3; 147/3  
C/Superfamily: annexin I; annexin repeat homology

Query Match 64.4%; Score 1037.5; DB 2; Length 316;  
Best Local Similarity 63.1%; Pred. No. 2.7e-60;  
Matches 200; Conservative 55; Mismatches 61; Indels 1; Gaps 1;

QY 1 MATLKVSDVPAAPSDAEQRTAFEGWGTNEDLIISILHRSABQRKVIROAVHETTYGD 60  
Db 1 MASLKVPATVPPEEDAEOLYKAFKGTNTERMIISILHARNATORSFIRAVVAAVYMKD 60  
QY 61 LRLTKDLKELSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMVEVACTRTSTQLH 120  
Db 61 LRLKELDEISGDERAVMLTFEPARDAYLAKESTMTFKNNVLEICTSALELFN 120  
QY 121 ARQAVHARYKKSLEEDVAHHTTDFRKLIVSVTSRYEGDEVNMTLAKQEAKVHEKIK 180  
Db 121 AKQAVQARYKTSLEEDVAHHTSGDIRKLIVPLVSTPRYDDEVNMTLASEAKILHEKIK 180  
QY 181 DKRYNDEVYRIILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALLRSTIOCL 240  
Db 181 EKAIVADDDILRIILSTRSKAQISATLNHYNNKNNFGTSMKSKYKE-DSENEYVQLKAKAVIKCL 239  
QY 241 TRRELAVFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIAGEYGRNSIPLEKAITKOT 300  
Db 240 TYREKTFEKVLROALINKLGTDEWGLTRVTTTRAEFDWERIKEEYIRNSVPLDRAIAKOT 299  
QY 301 RGDYERKMLVALLGEDD 317  
Db 300 HGDYEDILALLGHDA 316

## RESULT 9

annexin P33 - maize  
C/Species: Zea mays (maize)  
C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C/Accession: T02961  
R/Patney, N.H.; James, N.C.; Greenland, A.J.  
Plant. Physiol. 112, 1391-1396, 1996  
A/Title: CDNA isolation and gene expression of maize annexins P33 and P35.  
A/Reference number: Z14796; MUID:97092863; PMID:8938425  
A/Accession: T02961  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-314 <BAT>  
A/Cross-references: UNIPROT:Q43863; EMBL:X98244; PIDN:CAA66900.1  
A/Experimental source: cultivar clipper; root tip  
C/Superfamily: annexin I; annexin repeat homology  
P/14-85/Domain: annexin repeat homology <XNR>

Query Match 62.1%; Score 1000.5; DB 2; Length 314;  
Best Local Similarity 60.3%; Pred. No. 6.8e-58;  
Matches 190; Conservative 53; Mismatches 71; Indels 1; Gaps 1;

QY 1 MATLKVSDVPAAPSDAEQRTAFEGWGTNEDLIISILHRSABQRKVIROAVHETTYGD 60  
Db 1 MATLKVPATVPVPAVDCDQLRAKAFQGWGTNEALISILGHDAARALRAVAAEAYGSE 60  
QY 61 LRLTKDLKELSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMVEVACTRTSTQLH 120  
Db 61 LRLSTIDELSGFERAVIILMTLDPAREDAVLNEARKKPKGNRVLYEICTRTSAQIFA 120  
QY 121 ARQAVHARYKKSLEEDVAHHTTDFRKLIVSVTSRYEGDEVNMTLAKQEAKVHEKIK 180  
Db 121 TROAVHERPKRSLEEDIAHVTGDPFKLLVPLVSTRYGDPVNTRLHASEAKLHEKIH 180  
QY 181 DKRYNDEVYRIILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALLRSTIOCL 240  
Db 181 HRAVSDDELIRILTRSKFQQLATFNHYNDAGHRNKKQL-RADPODEVLTRLTRALIKCP 239  
QY 241 TRRELAVFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIAGEYGRNSIPLEKAITKOT 300  
Db 240 SCPDRYFEKVAQALAGLGTDENSLRTVITTRAEVDLKIKAQYGRNSVRLERAVAGDT 299  
QY 301 RGDYERKMLVALLGED 315  
Db 300 SGEYESMLALLGOE 314

RESULT 10  
T50026  
annexin-like protein - Arabidopsis thaliana  
N/Alternate names: protein T31P16.210  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004  
C/Accession: T50026  
R/Bevan, M.; Zimmermann, W.; Grueniseen, A.; Wambutt, R.; Kalicki, J.; Wohlmann, P.; St  
submitted to the Protein Sequence Database, May 2000  
A/Reference number: Z25027  
A/Accession: T50026  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-318 <BEV>  
A/Cross-references: UNIPROT:Q9LX08; EMBL:AL356332; GSPDB:GN00063; ATSP:T31P16.210  
A/Experimental source: cultivar Columbia; BAC clone T31P16  
C/Genetic:  
A/Gene: ATSP:T31P16.210  
A/Map position: 5  
A/Intons: 26/1; 74/3; 147/3  
C/Superfamily: annexin I; annexin repeat homology

Query Match 61.5%; Score 991.5; DB 2; Length 318;  
Best Local Similarity 60.8%; Pred. No. 2.6e-57;  
Matches 194; Conservative 52; Mismatches 70; Indels 3; Gaps 2;

QY 1 MATLKVSDVPAAPSDAEQRTAFEGWGTNEDLIISILHRSABQRKVIROAVHETTYGD 60

Db 1 MASLKIPANIPLEPDESEQLHAKFGWGTEGMIISILAHRNATQSFITAVVAANYNKD 60  
Qy 61 LKLTLDKELSNDFERRALLWTLEGGEDALLANATKWTSSNOVLMVACTRTSTOLH 120  
Db 61 LKLTLDKELSGDFERRVMTLTDPTERDAYLANESTKLPFKINVLVLEICTRPSLEFPK 120  
Qy 121 AROAYHARYKKSLEEDVAHHTTGPFRKLVSIVTSYVEEG--DEVNMTLAKOEKVLVHK 178  
Db 121 TKOAYHARYKTSLEEDVAHHTSGIRKLVPLVSTFYDGNADVKNKLARSEKTLHK 180  
Qy 179 IKOHVNDVEDVIRILSTRSKAQINATFNRYODDHGEELIKSLBEGDDDKFLALRSTIQ 238  
Db 181 ITEKAYVDEDLIRLITLRSKAQINATLNFPRKFGSSINFKKE--DSNDVYOLKTKAIK 239  
Qy 239 CLTRPELYFVDVLRSAINKTGTDEGALTRIVTTTAEIDLVKVGEEYORNSIPLEKAITK 238  
Db 240 CLTVPEKYPFKVLRRAINMGTEDEWALTRVTTTAEVLERIKELRIRNSVPLDRAIAN 239  
Qy 299 DTRGDYKMLVALLGEDA 317  
Db 300 DTSGDYKOMLALLGHDA 318

## RESULT 11

T02975  
annexin p35 - maize  
C/Species: Zea mays (maize)  
C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C/Accession: T02975  
R/Battery, N.H.; James, N.C.; Greenland, A.J.  
Plant Physiol. 112, 1391-1396, 1996  
A/Title: cDNA isolation and gene expression of maize annexins p33 and p35.  
A/Reference number: Z14796; MUID:97092863; PMID:8938425  
A/Accession: T02975  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-314 <BAT>  
A/Cross-references: UNIPROT:Q43864; EMBL:X98245; NID:91370602; PIDN:CAA6901.1; PID:g137  
A/Experimental source: cultivar clipper; root tip  
C/Superfamily: annexin I; annexin repeat homology  
F/14-85/Domain: annexin repeat homology <AXR>

Query Match 58.9%; Score 949.5; DB 2; Length 314;  
Best Local Similarity 58.7%; Pred. No. 1,4e-54;

Matches 185; Conservative 51; Mismatches 78; Indels 1; Gaps 1;

Qy 1 MATLKVSQSVPAAPSDAEOLRTAFEGWGTNEDLIISILAHSAEORKVIRQAYHETGCD 60  
Db 1 MATLTPSSVPAAVAEDCEQLHAKFEGWGTNEKLIISILAHRNAQAARIRRGYAEVYGE 60  
Qy 61 LKLTLDKELSNDFERRALLWTLEGGEDALLANATKWTSSNOVLMVACTRTSTOLH 120  
Db 61 LKLTLDKELSGDFERRVMTLTDPTERDAYLANESTKLPFKINVLVLEICTRPSLEFPK 120  
Qy 121 AROAYHARYKKSLEEDVAHHTTGPFRKLVSIVTSYVEEG--DEVNMTLAKOEKVLVHK 178  
Db 121 TKOAYHARYKTSLEEDVAHHTSGIRKLVPLVSTFYDGNADVKNKLARSEKTLHK 180  
Qy 179 IKOHVNDVEDVIRILSTRSKAQINATFNRYODDHGEELIKSLBEGDDDKFLALRSTIQ 238  
Db 181 ITEKAYVDEDLIRLITLRSKAQINATLNFPRKFGSSINFKKE--DSNDVYOLKTKAIK 239  
Qy 239 CLTRPELYFVDVLRSAINKTGTDEGALTRIVTTTAEIDLVKVGEEYORNSIPLEKAITK 238  
Db 240 CLTVPEKYPFKVLRRAINMGTEDEWALTRVTTTAEVLERIKELRIRNSVPLDRAIAN 239  
Qy 299 DTRGDYKMLVALLGEDA 317  
Db 300 DTSGDYKOMLALLGHDA 318

## RESULT 12

S56674  
annexin homolog R14 (clone R14) - garden strawberry (fragment)  
C/Species: Fragaria x ananassa (garden strawberry)  
C/Date: 10-Oct-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C/Accession: S56674  
R/Wikinson, J.Q.; Lanahan, M.B.; Conner, T.W.; Klee, H.J.  
Plant Mol. Biol. 27, 1097-1108, 1995  
A/Title: Identification of mRNAs with enhanced expression in ripening strawberry fruit.  
A/Reference number: S56673; MUID:95284359; PMID:7766892  
A/Accession: S56674  
A/Status: preliminary; nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 1-271 <WIL>  
A/Cross-references: UNIPROT:P51074; EMBL:U19941; NID:9643075; PIDN:AAA79922.1; PID:96431  
C/Superfamily: annexin I; annexin repeat homology  
F/42-113/Domain: annexin repeat homology <AX2>  
F/125-196/Domain: annexin repeat homology <AX3>  
F/200-271/Domain: annexin repeat homology <AX4>

Query Match 48.0%; Score 772.5; DB 2; Length 271;  
Best Local Similarity 55.3%; Pred. No. 3.5e-43;

Matches 151; Conservative 50; Mismatches 69; Indels 3; Gaps 2;

Qy 45 ORKVIROAYHETGEEDLTKLDELNDFERALLWTLEGGEDALLANATKWTSSNQ 104  
Db 1 ERKEIRAAIEQYLQYEDLKPSELSDGFEKAYVRRITLDPADRDVLANVAIKSTIDVYN 60  
Qy 105 VLMVACTRTSTQQLHARQAYHARYKKSLEEDVAHHTTGPFRKLVSIVTSYVEEGEVN 164  
Db 61 VVIRISCHSPBELLAVRAYQVLRGYSVEEDLAHTTGDRLTVALVTARYDGHEN 120  
Qy 165 MTLAKOEALVHEKIKODKHNDVDVIRIISTRSKAQINATFNRYODDHGEELIKS--LEEG 223  
Db 121 AKLANSEADILDAIDAKAFNHEIRIISTRSKQTLMATFMYCRDDQGISIKNLBEG 180  
Qy 224 DDDDKFLALRSTIQCLTRPELYFVDVLRSAINKTGTDEGALTRIVTTTAEIDLVKVGEE 283  
Db 181 AND--FOKALHTAIRCLNDPKKTFEYKVLNKAIRKGVTDDEALTRIVTTTAEIRDIKEY 238  
Qy 284 YORNSIPLEKAITKDTGRDYKMLVALLGEDD 316  
Db 239 YRKNSVPLDRAIANVAKOTSGDYKAFLLTLGKED 271

## RESULT 13

A84809

probable annexin [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: A84809

R/lyn, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.  
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, L.  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MUID:20083467; PMID:10617197

A/Accession: A84809

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-321 <STO>

A/Cross-references: UNIPROT:Q9ZVU7; GB:AE002093; NID:93785996; PIDN:AAC67342.1; GSPDB:GA

A/Gene: At2g38760

A/Map position: 2

Query Match 37.4%; Score 602.5; DB 2; Length 321;  
Best Local Similarity 39.0%; Pred. No. 4.7e-32;

Matches 124; Conservative 66; Mismatches 123; Indels 5; Gaps 1;

Qy 1 MATLKVSQSVPAAPSDAEOLRTAFEGWGTNEDLIISILAHSAEORKVIRQAYHETGCD 60  
Db 1 MATIRVPNEVPAPADSETLKQAIRGWTDEKAIIRVGLGRDQSGRKIRRESFRRIYKGD 60

OY	61	LKLTLDKLSNDPEFALILMTLEPGEERALLAN-----EATKRMTSSNOVLMEVACRITS	115
Db	61	LIDVLSSELSGDFEMKAVSWTYDDPAERDARLVNKLINKEKKKSULENLKIVELISCTTSP	120
OY	116	TOLLHQAQVARYARKYSLEEDVAHHHTTGDFRLTLVSLTSRYREGDEVMNTLAKOEAKTV	175
Db	121	NHLIAVRKAQVCSLPFSSLEEHIASLPPFLKLVITLASTRYDKORDAIDAVAITEAML	180
OY	176	HEKIKOKHYNEDEVIRIISTRSKAOINATEPNKYODDHCEBELKLSLBEGDDDDKFLALIRS	235
Db	181	REAEIEKKOLDHDHVLYIGTRSIYOLRETFAAYKNKYAVTIDKVGDGCPGADURLSLTKV	240
OY	236	TIQCITREPETYFVDVLAISAINKGTDEGALVTRIYTTTRAIEDLKAVIGEORRNISIPLEKA	295
Db	241	AIFCIDTBECHKPAKVYRDSIEGFIDEBSLTRAIYTRAEIDLMMKVRGSEFYMYMNTSMDNA	300
OY	296	ITKOTRGDYERKKVALLG	313
Db	301	ITGDISGDYKDFIITLLG	318

RESULT 14

996704  
996704  
probable annexin T23K23.6 [imported] - Arabidopsis thaliana  
C|Species: Arabidopsis thaliana (mouse-ear cress)  
C|Date: 02-Mar-2001 #sequence\_revision 02-Mar-2004 #text\_change 09-Jul-2004  
C|Accession: B96704  
R|Thelodopsis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A|Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzilli  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A|Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.W.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.  
A|Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A|Reference number: AB6141; MUID:21016719; PMID:11130712  
A|Accession: B96704  
A|Status: Preliminary  
A|Molecule type: DNA  
A|Residues: 1-316 <STO>  
C|Genetics: C|Genetics: UNIPROT:Q9C9X3; GB:AB005173; NID:95553887; PIDN:AAF16553.1; GSPDB:GN  
A|Gene: T23K23.6  
A|Map position: 1  
C|Superfamily: annexin I; annexin repeat homology

Query Match 32.5%; Score 523; DB 2; Length 316;  
Beet Local Similarity 34.7%; Pred. No. 6.7e-27;  
Matches 109; Conservative 75; Mismatches 128; Indels 2; Gaps 2;

QY 1 MATLKVSVPAPSDAEQLPTAFEGMGTNEDLISILAHRSAPBOKVIRQAYHETVEED 60  
1 MATMKIPIPTVSPSPRADAQOLPKAFKGTGCDTSVIINILAHNAQORALIEDEYETKPSDD 60  
DB 1 LKTLIDKELSNDFERAILLWLTLEPGERDALLANETKRWTSNQVLMEVACTRISTQILLH 120  
61 LKRLHSELHGLHKKAVLLMPPEAVERDASILKSLRGAVTDHKAIAEIICTRSQSQRQ 120  
DB 1 ARQAVHAQKKSLEEDVAHHTTGDFRKLIVLSVSYRGEDEVMTLAKQAKVHEKIK 180  
121 IKQVSNTEGVGLIEDISEASGNHKKVILAVLNTTRYEGEIDINASVENDARTLSKAVA 180  
QY 181 DKHVNDV-VRIILSTRKAQINATFNRYQDDHGEILKSLSEGGDDDKFLALLRSTQC 239  
DB 181 RKHNSDDDTLLQIFDRSRTHLVAVRSTYSNKGELKALRD-ETRGNFEHVLITTIQC 239  
QY 240 LTRPELVVDVYLSAINTKGTDEGALTRIIVTFRADIKVIGEYQRENISPLEKATKD 299  
DB 240 AENSFFYPAKALKRSMKGLGTDITLIRIVVTRAEVDNQFIITEYRKRYKKTILNVAHSD 299  
QY 300 TRGDYKMLVALLG 313

Db 300 TTSHYRTFLSLIG 313

## RESULT 15

LUHUIS  
annexin XIII, intestinal [validated] - human

C:\Species: Homo sapiens (man)  
C:\date: 31-Dec-1992 #sequence 31-Dec-1992 #text \_change 09-Jul-2004

C/Accession: A41/33/ 51016z  
R/Wice, B.M.; Gordon, J.I.  
T Cell Biol. 116, 405-422. 1992

A1 Title: A strategy for isolation of cDNAs encoding proteins affecting human intestinal A1 Reference number: A41733; MUID:92112982; PMID:1530946

A:Accession: A41733  
A:Molecule type: mRNA

A:Residues: 1-316 <MIC>  
A:Cross-references: UNIPROT:P27216; EMBL:Z11502; NID:g33919; PIDN:CAA7578.1; PID:g3398

C1: comment: Annexins undergo reversible, calcium-dependent binding to membrane phospholipids.

C/Genetics:  
A;Gene: GDB:ANXA13  
A;Cross-References: GDB:9836021. OMIM:602573

C:Superfamily: annexin I; annexin repeat homology  
A:Map position: 8q24.1-8q24.2  
M:100% identical CDS: ccdn:000022 | chr11:000070

**C-Keywords:** blocked amino end; calcium binding; duplication; endonexin fold; lipoprotein F2-316/**Product:** annexin intestine-specific #status experimental <MAT>

F:17-88/Domain: annexin repeat homology <AX1>  
F:28-44/Region: endonexin fold #status predicted

F:89-160/Domain: annexin repeat homology <AX2>  
F:100-116/Region: endonexin fold #status predicted

F:172-244/Domain: annexin repeat homology <AX>  
F:184-200/Region: endonexin fold #status predicted  
E:248-316/Domain: annexin repeat homology <AX>

F-2/Modified site: myristylated amino end (S1V) (in mature form) #status experimental

	F3/Modified	aspartic acid (Asn)	#status predicted
--	-------------	---------------------	-------------------

Query Match	31.1%;	Score 501;	DB 1;	Length 316;
Best Local Similarity	38.9%;	Pred. No. 1.8e-25;		

Matches	117;	Conservative	61;	Mismatches	115;	Indels	8;	Gaps	3;
---------	------	--------------	-----	------------	------	--------	----	------	----

[illegible]

76 AITIMTEEDGPRNIIANEA TRYBWSNOVI MEVA CRTSTOIIJAHOA VYHARYKSI EE 135

Db 79 TALALLDPPEYAROLQAMKSLGTDESVLIEFLCTRTINKEIILAKEYVQRLFDRSLLES 138

136 DVAAHTTGDRKLLSVLTVSYRVEGDEVNMTAKQKAKLVHEKIKDKHYNDEVVI - RILS 194

Db 139 DYAGDTSGLKKTLLVSLQANNEGDDVDKDLAQDAKDLVDAGEGRMGTEBELAFNEVLA 198

195 TRSKAQINAFENRYODDHGEELIKSLSE--GDDDDKFLALLRSTIQCTRPILYFVDVL 251

```

Db      199 KRKYKQLPATFQAYQILIGKIDIEAIEEETSGDLQKAYLTIVRCADD---EDYFAERL 254

```

252 R SAINKTGDEGALTRIVTTRAEIDLKVI GEFYQRNSIPLEKATTKDTRGDYEMVLVAL 311

Db 255 YKSMKAGTDEETLRIIVTTRAEVDLOGI KAKFQEKYQKSLSDWTRSDPSGDFPKLLVAL 314

QY 312 L 312

Db	315 L 315
----	-----------

Search completed: August 19, 2005, 20:47:17

Job time : 41 secs







```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6593
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6593
```

Query Match 29.3%; Score 472; DB 4; Length 323;

Best Local Similarity 36.3%; Pred. No. 5.6e-38; Matches 111; Conservative 63; Mismatches 130; Indels 2; Gaps 2;

```
Qy 12 APSDDAQLRTAFEGMGTNEDLIISILAHSAEQRYIROAYHETYGEDLLKTLDEKLSN 71
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 19 SPVDMAIQAIRGISTDEKMLISILTERSNARQILVKEYQAYKEKLDKDLGSLG 78
Qy 72 DFERAILMTLEPGERDALLANEATKRWTSNOVLMVACTRTSTOLLHARQAYHARYKK 131
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 79 HFEHLWALVTPPAVFAKQKKSMKGAGTNEIDLILTRTSRQMKDISQAYVTVYKK 138
Qy 132 SLEEDVAHHTTGDPRKLLVSLVTSYREGDEVNMTLAKQAKLVHEKIKKHYNDED-VI 190
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 139 SLDDDISSETSGDFRKLTLTLADGRDESLKVDENHLAKQDQILYKAGENRMGTDEDKFT 198
Qy 191 RIISTRSKAQINATFNRYODDHGEIILKSLSEGGDDDKFLALRSTIOCTLPRELYFVDV 250
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 199 EILCSFPQKLTFFEDYRNISQKDIYDSI-KGELSGHFEDLLALVNCVRNTPAFILAER 257
Qy 251 LRSAINKTGDEGALTRIVTTAEIDLKVI GEYQORNSIPLKATIKOTRGYEMKLV 310
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 258 LHRALKGIGTDEFTLRNIMVSRSEIDLDIRTEFKKHGYSLVSAIKSDTSGDYETITLK 317
Qy 311 LAGEDD 316
Db 318 ICGGDD 323
```

RESULT 6  
US-09-949-016-7952

```
; Sequence 7952, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7952
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7952
```

Query Match 29.3%; Score 472; DB 4; Length 334;

Best Local Similarity 36.3%; Pred. No. 5.9e-38; Matches 111; Conservative 63; Mismatches 130; Indels 2; Gaps 2;

```
Qy 12 APSDDAQLRTAFEGMGTNEDLIISILAHSAEQRYIROAYHETYGEDLLKTLDEKLSN 71
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 30 SPVDMAIQAIRGISTDEKMLISILTERSNARQILVKEYQAYKEKLDKDLGSLG 89
Qy 72 DFERAILMTLEPGERDALLANEATKRWTSNOVLMVACTRTSTOLLHARQAYHARYKK 131
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 90 HFEHLWALVTPPAVFAKQKKSMKGAGTNEIDLILTRTSRQMKDISQAYVTVYKK 149
```

```
Qy 132 SLEEDVAHHTTGDPRKLLVSLVTSYREGDEVNMTLAKQAKLVHEKIKKHYNDED-VI 190
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 150 SLDDDISSETSGDFRKLTLTLADGRDESLKVDENHLAKQDQILYKAGENRMGTDEDKFT 209
Qy 191 RIISTRSKAQINATFNRYODDHGEIILKSLSEGGDDDKFLALRSTIOCTLPRELYFVDV 250
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 210 EILCSFPQKLTFFEDYRNISQKDIYDSI-KGELSGHFEDLLALVNCVRNTPAFILAER 268
Qy 251 LRSAINKTGDEGALTRIVTTAEIDLKVI GEYQORNSIPLKATIKOTRGYEMKLV 310
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 269 LHRALKGIGTDEFTLRNIMVSRSEIDLDIRTEFKKHGYSLVSAIKSDTSGDYETITLK 328
Qy 311 LAGEDD 316
Db 329 ICGGDD 334
```

RESULT 7  
US-09-949-016-10475

```
; Sequence 10475, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10475
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10475
```

Query Match 28.6%; Score 461.5; DB 4; Length 681;

Best Local Similarity 35.7%; Pred. No. 1.8e-36; Matches 112; Conservative 58; Mismatches 129; Indels 15; Gaps 4;

```
Qy 13 PSDDAQLRTAFEGMGTNEDLIISILAHSAEQRYIROAYHETYGEDLLKTLDEKLSND 72
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 373 PDADAALAKKMGAGTDEDTIIDITTHSNVORQIROTFSKHFERDMLTKSEISGD 432
Qy 73 FERRAILMTLEPGERDALLANEATKRWTSNOVLMVACTRTSTOLLHARQAYHARYKS 132
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 433 LARLILGLMMPRAHYDAKQKKAMGAGTDEKALILATRTVAIRAINAEVKEDYHKS 492
Qy 133 LSEEDVAHHTTGDPRKLLVSLVTSYREGDEVNMTLAKQAKLVHEKIKK-----DKHN 185
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 493 LEDALSSDPSGHRRLILSLATGHRERGE-NLDQAREDAQVAATLEADTPSSGKTSI 551
Qy 186 DEDVIRIISTRSKAQINATFNRY--QDDHGEIILKSLSEGGDDDKFLALRSTIOCTLR 242
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 552 ETRFMITLCTRSPHARRVFOEIRKMTNVDEHTIKKMSGVDRDAFVAI-----VQSVKN 607
Qy 243 PELYFVDVLSAINKTGDGALTRIVTTAEIDLKVI GEYQORNSIPLKATIKOTRG 302
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 608 KPLFPADKLYKSMKGAGTDEKTLTRIMVSRSEIDLINRREFLEKDKSLHQALBEDTSG 667
Qy 303 DYKMTVALGDED 316
Db 668 DFLKALLCGGED 681
```

RESULT 8  
US-09-949-016-10476

```
/ Sequence 10476, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 10476
/ LENGTH: 681
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-016-10476

Query Match      28.6%; Score 461.5; DB 4; Length 681;
Best Local Similarity 35.7%; Pred. No. 1.8e-36;
Matches 112; Conservative 58; Mismatches 129; Indels 15; Gaps 4;

QY      13  PSDDAEQRTAFEGGTGNTDILISLAHRSARQKVIROAVHETYGEDLLKTLDEKLSND 72
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      373  PDADAALRKAMGGLGTDEBTITIDITHRSNVORQOIRQTFKSHFGRDLTKSEISGD 432

QY      73  FERAILLWLEPGERDALANEATKRWTSNQVLMVACTRTSTOLHARQAVHARYKS 132
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      433  LARLILGLMPPRAHYADKQKXMEGAGTDEKALILILATRAAEIRAINAEKEDYHKS 492

QY      133  LEEDVAHHTGDPFRKLIVSLVTSYRYEGDEVNMTLAKQBAKLVHEKIK-----DKHYN 185
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      493  LEDALSDTSGHFRRLISLATGHRREGGE-NDQARBDAAQVAETILEIADTPSGDKTSL 551

QY      186  DEDVIRILSTRSAQINATFNRY---QDDHGEILSLBEGDDDDFLLARSTQCLR 242
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      552  ETRFMWILCTRSYPHLRVFOEIRKMTNVDEHTIKKMSGVDRDAFVAI---VQSVKN 607

QY      243  PELYPDVLRSAINTGTDEGALTRIVTTTAAIDLVIGEYORRNSIPLEKAITKDTGTG 302
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      608  KPFPADKLYKSKKAGTDEKTLTRIMVSRSEIDLINIRREFLEKDKSLHQAIIEGDTSG 667

QY      303  DYKXMLVALLGEDD 316
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      668  DFLKALLALCGED 681

RESULT 9
US-09-949-016-7621
/ Sequence 7621, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7621
/ LENGTH: 557
```

```
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-016-7621

Query Match      28.0%; Score 451; DB 4; Length 557;
Best Local Similarity 35.7%; Pred. No. 1.4e-35;
Matches 110; Conservative 61; Mismatches 129; Indels 8; Gaps 3;

QY      13  PSDDAEQRTAFEGGTGNTDILISLAHRSARQKVIROAVHETYGEDLLKTLDEKLSND 72
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      254  PLRDAEVLKAMKGGEIDQALIIDCLGRSNKQROQILSFRTAYGKDLIKOLKSELSEN 313

QY      73  FERAILLWLEPGERDALANEATKRWTSNQVLMVACTRTSTOLHARQAVHARYKS 132
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      314  FEKTILAKMKTVPDPIYEIKALINGVDEACILIELLASRNEHIRELNRAYKAEFKKT 373

QY      133  LEEDVAHHTGDPFRKLIVSLVTSYRYEGDEVNMTLAKQBAKLVHEKIKDKHYNDEYIR- 191
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      374  LEEAIRSDTSGHFORLLISLQGNFDESTYVMISLAQRDAQCELYAAGEVRLGTDSKFNA 433

QY      192  ILSTRSKQINATFNRYODDHEEILKSL---EBGDDDDKFLALLRSTIQCLTRPELVY 248
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      434  VLCSSRAHLVAVFENYQRTGRDIEKSTCREMSGDLEGMLAV---VKCLAKTTPAFPA 489

QY      249  DVLRSAINTGTDEGALTRIVTTTAAIDLVIGEYORRNSIPLEKAITKDTGRDYEXML 308
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      490  ERLNKAAMRGAGTKDILIRIMVSRSETLDIRSEYKKNYKSKLVHDISGDTSGDYRKIL 549

QY      309  VALLGEDD 316
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB      550  LKICGGND 557

RESULT 10
US-08-526-136-2
/ Sequence 2, Application US/08526136
/ Patent No. 6107089
/ GENERAL INFORMATION:
/ APPLICANT: Towle, Christine A. et al.
/ TITLE OF INVENTION: ANNEXIN XI
/ NUMBER OF SEQUENCES: 36
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: U.S.A.
/ ZIP: 02110-2804
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ COMPUTER: IBM PS/2 Model 502 or 55SX
/ OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
/ SOFTWARE: WordPerfect (Version 5.0)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/526,136
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/214,036
/ FILING DATE:
/ APPLICATION NUMBER: 07/837,775
/ FILING DATE: February 13, 1992
/ APPLICATION NUMBER: 07/764,465
/ FILING DATE: September 23, 1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Clark, Paul T.
/ REGISTRATION NUMBER: 30,162
/ REFERENCE/DOCKET NUMBER: 00786/099001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 542-5070
/ TELEFAX: (617) 542-8906
/ TELEX: 200154
/ INFORMATION FOR SEQ ID NO: 2:
```



```
/ Patent No. 6107089
/ GENERAL INFORMATION:
/ APPLICANT: Towle, Christine A. et al.
/ TITLE OF INVENTION: ANNEXIN XI
/ NUMBER OF SEQUENCES: 36
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: U.S.A.
/ ZIP: 02110-2804
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ COMPUTER: IBM PS/2 Model 50z or 55sx
/ OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
/ SOFTWARE: WordPerfect (Version 5.0)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/526,136
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/214,036
/ FILING DATE:
/ APPLICATION NUMBER: 07/837,775
/ FILING DATE: February 13, 1992
/ APPLICATION NUMBER: 07/764,465
/ FILING DATE: September 23, 1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Clark, Paul T.
/ REGISTRATION NUMBER: 30,162
/ REFERENCE/DOCKET NUMBER: 00786/099001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 542-5070
/ TELEFAX: (617) 542-8906
/ TELEX: 200154
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 466
/ TYPE: amino acid
/ STRANDEDNESS: N/A
/ TOPOLOGY: N/A
/ US-08-526-136-13

Query Match      27.7%; Score 447; DB 3; Length 466;
Best Local Similarity 33.1%; Pred. No. 2,7e-35;
Matches 102; Conservative 67; Mismatches 121; Indels 18; Gaps 3;

QY 16 DAEQLRTAFEGWGTNEDLIISLAHRSABQRKVIROAYHETYGEDLLKTLDEKLSNDFR 75
DB 168 DAEILRKAMKGFSTDEQAIYDVANRSDQRKIKAAFKTSYKGLIKDLKSELSGNMEE 227
QY 76 AILLMTLEPERGALLANEATKRWTSNQVLEMAVCTSTQQLHARQAYHARKYKSLSE 135
DB 228 LILALFMPPTYYDAWSLRKAMQAGTQERVLIELCTRNQELREIVRCYQSEFGRDLK 287
QY 136 DVAAHTTGFRLKLVSLVTSYREGDEVNMTLAKQEAKEIKDKHYNDEVIR-ILS 194
DB 288 DIRSDTSGHFERLLVSMQGNRDNOSINHQAQSDAQLVYAGSGRLGTDESCNMILA 347
QY 195 TRSKAQINATFNRYODDHEEILKS-----LEGGDDDKFLALLSTIOCLTRPELY 246
DB 348 TRSFPQLRATMEAYSRMANRDLSSVSRSFSGYVSG-----LKTILQCALNRPAF 398
QY 247 FVDVRSAINKTGTDEGALTRIVTTRARIDLVIGEYQRRNSIPLEKAITDTRGDYK 306
DB 399 FAERLYYAKKAGAGTDDSTLVRIVTRSEIDLVOIKQMFQMYOKTLGTMIAGTSGDYR 458
QY 307 MLVALLGE 314
DB 459 LLLAIVGQ 466
```

```
RESULT 14
US-09-949-016-7070
/ Sequence 7070, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 7070
/ LENGTH: 467
/ TYPE: PRT
/ ORGANISM: Human
/ US-09-949-016-7070
```

Query Match 27.7%; Score 447; DB 4; Length 467;

Best Local Similarity 33.1%; Pred. No. 2,7e-35;

Matches 102; Conservative 67; Mismatches 121; Indels 18; Gaps 3;

```
QY 16 DAEQLRTAFEGWGTNEDLIISLAHRSABQRKVIROAYHETYGEDLLKTLDEKLSNDFR 75
DB 169 DAEILRKAMKGFSTDEQAIYDVANRSDQRKIKAAFKTSYKGLIKDLKSELSGNMEE 228
QY 76 AILLMTLEPERGALLANEATKRWTSNQVLEMAVCTSTQQLHARQAYHARKYKSLSE 135
DB 229 LILALFMPPTYYDAWSLRKAMQAGTQERVLIELCTRNQELREIVRCYQSEFGRDLK 288
QY 136 DVAAHTTGFRLKLVSLVTSYREGDEVNMTLAKQEAKEIKDKHYNDEVIR-ILS 194
DB 289 DIRSDTSGHFERLLVSMQGNRDNOSINHQAQSDAQLVYAGSGRLGTDESCNMILA 348
QY 195 TRSKAQINATFNRYODDHEEILKS-----LEGGDDDKFLALLSTIOCLTRPELY 246
DB 349 TRSFPQLRATMEAYSRMANRDLSSVSRSFSGYVSG-----LKTILQCALNRPAF 399
QY 247 FVDVRSAINKTGTDEGALTRIVTTRARIDLVIGEYQRRNSIPLEKAITDTRGDYK 306
DB 400 FAERLYYAKKAGAGTDDSTLVRIVTRSEIDLVOIKQMFQMYOKTLGTMIAGTSGDYR 459
QY 307 MLVALLGE 314
DB 460 LLLAIVGQ 467
```

RESULT 15

US-09-949-016-7071

/ Sequence 7071, Application US/09949016

/ Patent No. 6812339

/ GENERAL INFORMATION:

/ APPLICANT: VENTER, J. Craig et al.

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

/ FILE REFERENCE: CLO01307

/ CURRENT APPLICATION NUMBER: US/09/949,016

/ CURRENT FILING DATE: 2000-04-14

/ PRIOR APPLICATION NUMBER: 60/241,755

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR APPLICATION NUMBER: 60/237,768

/ PRIOR FILING DATE: 2000-10-03

/ PRIOR APPLICATION NUMBER: 60/231,498

/ PRIOR FILING DATE: 2000-09-08

/ NUMBER OF SEQ ID NOS: 207012

/ SOFTWARE: FASTSEQ for Windows Version 4.0



This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 20:27:39 ; Search time 161 Seconds

(without alignments)  
771.011 Million cell updates/sec

Title: US-10-690-564-2

Sequence: 1 MATLKXSDSVAPSDAEOL.....KOTRGDYKLVALLGEDDA 317

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubppa/PCUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubppa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/2/pubppa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/2/pubppa/US11\_PUBCOMB.pep:\*  
20: /cgn2\_6/ptodata/2/pubppa/US11\_NEW\_PUB.pep:\*  
21: /cgn2\_6/ptodata/2/pubppa/US60\_NEW\_PUB.pep:\*  
22: /cgn2\_6/ptodata/2/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1611	100.0	317	US-10-690-564-2	Sequence 2, Appl
2	1185.5	73.6	316	US-10-393-840-56	Sequence 56, Appl
3	1177.5	73.1	313	US-10-424-599-144032	Sequence 144032,
4	1177.5	73.1	333	US-10-425-114-47663	Sequence 47663, A
5	1177.5	73.1	333	US-10-425-114-55802	Sequence 55802, A
6	1171.5	72.7	316	US-10-219-220-62	Sequence 62, Appl
7	1171.5	72.7	316	US-10-393-840-114	Sequence 114, Appl
8	1132.5	70.3	316	US-10-219-220-260	Sequence 260, Appl
9	1054.5	65.5	314	US-10-424-599-213299	Sequence 213299,
10	1011.5	62.8	314	US-10-767-701-45089	Sequence 45089, A
11	1009.5	62.7	339	US-10-425-114-60631	Sequence 60631, A

12	1009.5	62.7	340	US-10-425-115-186841	Sequence 186841,
13	1009.5	62.7	361	US-10-425-114-64727	Sequence 64727, A
14	1008.5	62.6	341	US-10-425-114-60513	Sequence 60513, A
15	1008.5	62.6	344	US-10-425-114-59289	Sequence 59289, A
16	1004.5	62.4	317	US-10-437-963-180790	Sequence 180790,
17	988.5	61.4	333	US-10-425-114-40052	Sequence 40052, A
18	977	60.6	315	US-10-437-963-161226	Sequence 161226,
19	975.5	60.6	314	US-10-767-701-46109	Sequence 46109, A
20	951.5	59.1	314	US-10-425-115-248356	Sequence 248356,
21	949.5	58.9	314	US-10-739-930-7820	Sequence 7820, Ap
22	919.5	57.1	245	US-10-425-114-43123	Sequence 43123, A
23	866.5	53.8	239	US-10-219-220-65	Sequence 65, Appl
24	866.5	53.8	239	US-10-393-840-54	Sequence 54, Appl
25	866.5	53.8	239	US-10-393-840-117	Sequence 117, Appl
26	827.5	51.4	304	US-10-424-599-231427	Sequence 231427,
27	823	51.1	312	US-10-424-599-172904	Sequence 172904,
28	799.5	49.6	313	US-10-424-599-230318	Sequence 230318,
29	786.5	48.8	313	US-10-424-599-279849	Sequence 279849,
30	786.5	48.8	320	US-10-425-114-55270	Sequence 55270, A
31	786.5	48.8	324	US-10-425-114-43962	Sequence 43962, A
32	782	48.5	314	US-10-739-930-9279	Sequence 9279, Ap
33	776	48.2	314	US-10-424-599-279847	Sequence 279847,
34	770	47.8	270	US-10-424-599-211315	Sequence 211315,
35	759	47.1	314	US-10-424-599-273745	Sequence 273745,
36	721	44.8	242	US-10-393-840-55	Sequence 55, Appl
37	691.5	42.9	289	US-10-437-963-121221	Sequence 121221,
38	635.5	39.4	241	US-10-425-114-41196	Sequence 41196, A
39	635.5	39.4	369	US-10-437-963-169438	Sequence 169438,
40	629	39.0	373	US-10-425-115-213961	Sequence 213961,
41	621.5	38.6	489	US-10-437-963-169441	Sequence 169441,
42	613.5	38.1	320	US-10-424-599-201908	Sequence 201908,
43	573	35.6	184	US-10-219-220-66	Sequence 66, Appl
44	573	35.6	184	US-10-393-840-118	Sequence 118, Appl
45	563	34.9	184	US-10-393-840-52	Sequence 52, Appl

#### ALIGNMENTS

RESULT 1  
US-10-690-564-2  
; Sequence 2, Appl US/10690564  
; Publication No. US20050089872A1  
; GENERAL INFORMATION:  
; APPLICANT: KOREA KIMHO PETROCHEMICAL CO., LTD  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ANNEXINS FROM PLANTS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/690,564  
; CURRENT FILING DATE: 2003-10-23  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Kopatentlin 1.71  
; SEQ ID NO 2  
; LENGTH: 317  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-690-564-2

Query Match 100.0%; Score 1611; DB 17; Length 317;  
Best Local Similarity 100.0%; Pred. No. 1.5e-130;  
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1 MATLKXSDSVAPSDAEOLRTAFEGKGTNEDELIISLAHSAEQRKVIROYHETGSD 60
DB	1 MATLKXSDSVAPSDAEOLRTAFEGKGTNEDELIISLAHSAEQRKVIROYHETGSD 60
QY	61 LKTKLDKELNSDFERAILLTLEPERDALANETKXTWTSNOVLMEVACTRTSTOLH 120
DB	61 LKTKLDKELNSDFERAILLTLEPERDALANETKXTWTSNOVLMEVACTRTSTOLH 120
QY	121 AROAHYARYKSLIEDVAHHTTGDFRKLVLVTSYRREGDEVNMTLAKQBAKVHEKIX 180
DB	121 AROAHYARYKSLIEDVAHHTTGDFRKLVLVTSYRREGDEVNMTLAKQBAKVHEKIX 180



```
QY      181 DKRYNDEVYRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALIRSTIOQL 240
      |||||
      181 DKRYNDEVYRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALIRSTIOQL 240
DB      241 TRBELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIGERYQRNSIPILEKAITKOT 300
      |||||
      241 TRBELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIGERYQRNSIPILEKAITKOT 300
DB      241 TRBELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIGERYQRNSIPILEKAITKOT 300
      |||||
      241 TRBELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIGERYQRNSIPILEKAITKOT 300
QY      301 RGDYKMLVALIGEDDA 317
      |||||
      301 RGDYKMLVALIGEDDA 317
DB      301 RGDYKMLVALIGEDDA 317
      |||||
      301 RGDYKMLVALIGEDDA 317
```

```
RESULT 2
US-10-393-840-56
; Sequence 56, Application US/10393840
; Publication No. US2003022922A1
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1012c3
; CURRENT APPLICATION NUMBER: US/10/393,840
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Esulalyptus grandis
US-10-393-840-56
```

```
Query Match      73.6%; Score 1185.5; DB 15; Length 316;
Best Local Similarity 71.9%; Pred. No. 7.8e-94;
Matches 228; Conservative 38; Mismatches 50; Indels 1; Gaps 1;

QY      1 MATLKYSDSVPAPSDAEQLRTAFEGWGNTNEDLIISILAHSAEQKVIROAYHETTYGD 60
      |||||
      1 MATLKYSDSVPAPSDAEQLRTAFEGWGNTNEDLIISILAHSAEQKVIROAYHETTYGD 60
DB      1 MSTLTVPQPLPVPVADDCQQLRTAFAGWGNTNEDLIISILAHSAEQKVIROAYHETTYGD 60
      |||||
      1 MSTLTVPQPLPVPVADDCQQLRTAFAGWGNTNEDLIISILAHSAEQKVIROAYHETTYGD 60
QY      61 LKTLDKELSNDFERAIIIMTLTEPGERDALLANEATKRWTSNOVLMEVACTRTSTOLLH 120
      |||||
      61 LKTLDKELSNDFERAIIIMTLTEPGERDALLANEATKRWTSNOVLMEVACTRTSTOLLH 120
DB      61 LKTLDKELSNDFERAIIIMTLTEPGERDALLANEATKRWTSNOVLMEVACTRTSTOLLH 120
      |||||
      61 LKTLDKELSNDFERAIIIMTLTEPGERDALLANEATKRWTSNOVLMEVACTRTSTOLLH 120
QY      121 ARQAVHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAKLVHEKIK 180
      |||||
      121 ARQAVHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAKLVHEKIK 180
DB      121 ARQAVHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAKLVHEKIK 180
      |||||
      121 ARQAVHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAKLVHEKIK 180
QY      181 DKRYNDEVYRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALIRSTIOQL 240
      |||||
      181 DKRYNDEVYRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALIRSTIOQL 240
DB      181 DKRYNDEVYRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALIRSTIOQL 240
      |||||
      181 DKRYNDEVYRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALIRSTIOQL 240
QY      241 TRBELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIGERYQRNSIPILEKAITKOT 300
      |||||
      241 TRBELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIGERYQRNSIPILEKAITKOT 300
DB      241 TRBELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIGERYQRNSIPILEKAITKOT 300
      |||||
      241 TRBELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIGERYQRNSIPILEKAITKOT 300
QY      301 RGDYKMLVALIGEDDA 317
      |||||
      301 RGDYKMLVALIGEDDA 317
DB      301 RGDYKMLVALIGEDDA 317
      |||||
      301 RGDYKMLVALIGEDDA 317
```

```
RESULT 3
US-10-424-599-144032
; Sequence 144032, Application US/10424599
; Publication No. US20040031072A1
```

```
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 144032
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101073C.1.pep
US-10-424-599-144032
```

```
Query Match      73.1%; Score 1177.5; DB 15; Length 316;
Best Local Similarity 71.3%; Pred. No. 3.8e-93;
Matches 226; Conservative 42; Mismatches 48; Indels 1; Gaps 1;

QY      1 MATLKYSDSVPAPSDAEQLRTAFEGWGNTNEDLIISILAHSAEQKVIROAYHETTYGD 60
      |||||
      1 MATLKYSDSVPAPSDAEQLRTAFEGWGNTNEDLIISILAHSAEQKVIROAYHETTYGD 60
DB      1 MATLKYSDSVPAPSDAEQLRTAFEGWGNTNEDLIISILAHSAEQKVIROAYHETTYGD 60
      |||||
      1 MATLKYSDSVPAPSDAEQLRTAFEGWGNTNEDLIISILAHSAEQKVIROAYHETTYGD 60
QY      61 LKTLDKELSNDFERAIIIMTLTEPGERDALLANEATKRWTSNOVLMEVACTRTSTOLLH 120
      |||||
      61 LKTLDKELSNDFERAIIIMTLTEPGERDALLANEATKRWTSNOVLMEVACTRTSTOLLH 120
DB      61 LKTLDKELSNDFERAIIIMTLTEPGERDALLANEATKRWTSNOVLMEVACTRTSTOLLH 120
      |||||
      61 LKTLDKELSNDFERAIIIMTLTEPGERDALLANEATKRWTSNOVLMEVACTRTSTOLLH 120
QY      121 ARQAVHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAKLVHEKIK 180
      |||||
      121 ARQAVHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAKLVHEKIK 180
DB      121 ARQAVHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAKLVHEKIK 180
      |||||
      121 ARQAVHARYKKSLEEDVAHHTTGDPRKLVSLVTSYRYEGDEVNMTLAKOEAKLVHEKIK 180
QY      181 DKRYNDEVYRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALIRSTIOQL 240
      |||||
      181 DKRYNDEVYRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALIRSTIOQL 240
DB      181 DKRYNDEVYRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALIRSTIOQL 240
      |||||
      181 DKRYNDEVYRILSTRSKAQINATFNRYODDHGEELIKSLSEGGDDDKFLALIRSTIOQL 240
QY      241 TRBELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIGERYQRNSIPILEKAITKOT 300
      |||||
      241 TRBELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIGERYQRNSIPILEKAITKOT 300
DB      241 TRBELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIGERYQRNSIPILEKAITKOT 300
      |||||
      241 TRBELYFVDVLRSAINKGTDEGALTRIVTTTAEIDLKVIGERYQRNSIPILEKAITKOT 300
QY      301 RGDYKMLVALIGEDDA 317
      |||||
      301 RGDYKMLVALIGEDDA 317
DB      301 RGDYKMLVALIGEDDA 317
      |||||
      301 RGDYKMLVALIGEDDA 317
```

```
RESULT 4
US-10-425-114-47663
; Sequence 47663, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B
; APPLICANT: Tabaska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47663
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700661911_FLI.pep
US-10-425-114-47663
```

Query Match	73.1%;	Score 1177.5;	DB 15;	Length 333;
Best Local Similarity	71.3%;	Pred. No. 4.1e-93;		
Match 226;	Conservative 42;	Mismatches 48;	Indels 1;	Gaps 1;
QY	1	MATLKVSDBVPAPSDAEQLRTAFEGMGNNEDIISILAHRSARQKRVIRQAHEHTYGED	60	
		:		
Db	18	MATLKVPQPLPVPVADDCEDQKRAKFSMGNGNEELIVSILAHRNAQKRLRETYAQTYGED	77	
QY	61	LLKTLDKELSDNFERAILLMTLEPGERDALLANEATKRWTSNQVLMVACTRSTQTLH	120	
		:		
Db	78	LLKALDKELTSDFERLVHMLTLDASERDAFLANEATKRWTSNQVLEIACRSTSEQLFA	137	
QY	121	ARQATHAARKKSLEEBVAHHTTGDPFKLLVSLVTSFRYEGDEPNMTLAKOEAALVHEKIK	180	
		:		
Db	138	ARKAHVLYKKSLEEBVAHHTTGDPFKLLPLVSSRYRDEDEVNLTAKTEALTLHEKIS	197	
QY	181	DKHYNDEDEVYRILSTSRKAQINATFRYODDHEELIKSLSECDDDDKFLALRSTQCL	240	
		:		
Db	198	NKAYVDDDPTRILATRSRQINNTLNMHYKDAFGQDINKOL-KADPDEREISLRAIVKCL	256	
QY	241	TRPELVFVDVLSAINKTGTDEGALTRIVYTTRAEIDLKLYGEEYQRNRSIPLKCAITKOT	300	
		:		
Db	257	IRPEKCFEKVAVLAINKGTDEGALRRVAVTRADEVLDKNIADRYQRSSVPLERAIVKOT	316	
QY	301	RGDYERKMLVALLGEDDA	317	
		:		
Db	317	TGDYERKMLVALLGHDDA	333	

```

RESULT 5
US-10-425-114-55802
| Sequence 55802, Application US/10425114
| Publication No. US2004003488A1
| GENERAL INFORMATION:
| APPLICANT: Liu, Jingdong
| APPLICANT: Zhou, Yihua
| APPLICANT: Kovalic, David K.
| APPLICANT: Screen, Steven E
| APPLICANT: Tabaska, Jack E
| APPLICANT: Cao, Yongwei
| TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
| TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
| FILE REFERENCE: 38-21(53313)B
| CURRENT APPLICATION NUMBER: US/10/425,114
| CURRENT FILING DATE: 2003-04-28
| NUMBER OF SEQ ID NOS: 73128
| SEQ ID NO 55802
| LENGTH: 333
| TYPE: PRT
| ORGANISM: Glycine max
| FEATURE:
| OTHER INFORMATION: Clone ID: UC-GMROPIC033B01_FL1.pep
US-10-425-114-55802

Query Match      73.1%; Score 1177.5; DB 15; Length 333;
Best Local Similarity 71.3%; Pred. 4.1e-93;
Matches 226; Conservative 42; Mismatches 48; Indels 1; Gaps 1;

Oy      1  MATLVSDSVPAQDDABQLRTAFEGWGNEDLTISILHRSASEOKVIRQAAHYETGYED 60
Db      18  MATLVKVPQPLPVPVADDCBQLKAFSGMGNEELIVSILHHRNAQKKLIRETYAQTGYED 77
Oy      61  LKLTLDKELISNDFERAILLMTLEPGERDALLANEATKRWTSNQVIMEVACTRTSTOLH 120
Db      78  LKKALDKELISDFEPLVHMLMTLDSAEERDAFLANEATKKMTSSNQVLEIACTRSSQLPA 137
Oy     121  AROAHYARKKSLSEEDVAHHTTGDFPKLIVSLVTSVRYSGDEFNMFLAQOEAQLVHEKIR 180
Db     138  ARKAHVLKYSLEEDVAHHTTGDRKKLILPLVSSVRYSGDEVNLTALATKALKLEKIS 197
Oy     181  DKHYNDEDVIRILSTRSKAQINATFNRRYODDHGEETLKSLEBGDDDDKFLALLRSTIOCL 240
Db     198  NKAYNDDDFIRILARRSRQAINATLNHYDAAGODINKDL-KADPQDEFSLRLRATVKCL 256

```

Oy	24	TREKLFVNVLSAINIKGTDEGALTRVTVAETDLKVIGEYORANSIPLERAIKOT	308
		:::::::::::::::::::::	
Dd	257	IREBKFEKYRLAIRKGSTDGALTRVVAITRAEVDLKNIDETOSSVPEIRAIKOT	316
Oy	301	RGYEKMLVALLGEDDA	317
Dd	317	TGDYEKMLVALLGHDDA	333

```

RESULT 6
US-10-219-220-62
; Sequence 62, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; TITLE OF INVENTION: death and their use in the modification of plant development
; FILE REFERENCE: 11000.1022c1
; CURRENT APPLICATION NUMBER: US/10/219,220
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 62
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-219-220-62

```

[illegible]

```

, , PRIOR FILING DATE: 2000-08-10
, , PRIOR APPLICATION NUMBER: US 09/170,862
, , PRIOR FILING DATE: 1998-10-13
, , PRIOR APPLICATION NUMBER: US 60/148,426
, , PRIOR FILING DATE: 1999-08-11
, , PRIOR APPLICATION NUMBER: PCT NZ/99/00169
, , PRIOR FILING DATE: 1999-10-08
, , NUMBER OF SEQ ID NOS: 956
, , SOFTWARE: FastSeq For Windows Version 4.0
, , SEQ ID NO 114
, , LENGTH: 316
, , TYPE: prt
, , ORGANISM: Pinus radiata
US-10-393-840-114

```

Query Match	72.7%;	Score 1171.5;	DB 15;	Length 316;
Best Local Similarity	71.3%;	Pred. No. 1.3e-92;		
Matches 226; Conservative	39;	Mismatches 51;	Indels 1;	Gaps 1

QY 1 MATLKVS DSVAPAPSDAEQLRTAFEGWGNTEDLTISILARSAEQKRVIFQAHYHETGYED 60

Db 1 MSTLTVPQPLRPVADDCCEQLRTAFEGWGNTKELTISILGHNNAAQRKLHQYIAETGYED 60

QY	Db
61	61
LLKRLDKELSNDFERAILMTLTPBGRDALIANEATKRTWSSNQVLMVEVACTRTSTQLH	LLKALDRELINDFERLTVVMSLDPERDAVIANEATKRTWSSNQVLMIEIKTRSPQLM
120	120

QY 121 ARQAVHARYKSLIEDVAHHHTTGDPRKLVSIVTSRYEEDENVNNTLAKQEAKLVEHKIR 180

Db 121 ARQAVHARYKSMEDVAHHHTTGDPRKLVLPLGSSRYNDQEDENVNNTLAKAEKALIEHKIS 180

QY 181 DKHYNDEDVRLIRLSRKAQINATNRNQDHDGEEILKSLBEGDDDKFLALLRRSTIQCL 240  
:  
Db 181 EKAGHEDLIRLIRATRSKQAVNATLNHYKNFNGNINKDKLT-DPKDAFLTLIRATVKCL 239

QY 241 TRPELYFVNLRSAINKTGDEGALTRIVTRAEDIDKVGEEYGRNSTPLEKATTKDT 300  
 Db 240 TRPEKYFEKLRLLAINKRGDEGALTRIVTRAEDVDMKFTSESYGRNSTLPLDRAIVKDT 299

QY	301	RGDYERKMLVALLGEDDA	317
		: : :	
		: : :	
		:	
Db	300	TGDYERKMLLALIGHVEA	316

RESULT 8  
US-10-219-220-260

```

; Publication No. US2003008272A1
;
; GENERAL INFORMATION:
; APPLICANT: Flynn, Barry

```

FILE REFERENCE	TITLE OF INVENTION	TITLE OF INVENTION	TITLE OF INVENTION
11000.1022c1	Compositions affecting programmed cell death and their use in the modification of plant development	Compositions affecting programmed cell death and their use in the modification of plant development	Compositions affecting programmed cell death and their use in the modification of plant development

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```

; SOFTWARE: Fastseq for Windows Version 3.0
;
; SEO ID NO 260
; LENGTH: 316
;

```

i- ORGANISM: *Eucalyptus grandis*  
US-10-219-220-260

Best Local Similarity	68.8%;	Pred. No. 3e-89;	
Matches	218;	Conservative	46;
		Mismatches	50;
		Indels	1;
		Gaps	1

Db

1 MATIATVPSPSPSPAEADQLQKAFACMGITNEDLLISLAAHNNAGAKRKIVIRQTAYAEIYGED 60

QY	61	LKTLKDLKSLDNDFEPAIILMTLEPGEERDALLANEKTXKXTSSNOVLMEVACRTSTQLLH	120
Db	61	LLKALDKELSSDFEESVILLMTLDPKERDAPFASNEATKTLTSSNWMLMETACRSSNLELM	120
QY	121	ARQAHAYKYSLEEDVAHHTTGDPRKLLVSLVTSYRYEGEDVNMTLAKQEAKLVHEKIK	180
Db	121	VRQAHAYKYSLEEDIAYHTTGDPRKLLVPLASFRRYEGEPVNMTLARSEAKILHEKIKH	180
QY	181	DKHVNDEVIRILSTRSAQINAFNNRYQDDHGEIILKSLEEGDDDKFLALLRSTIOCL	240
Db	181	EKAYNHDELIRIVTTRSAQINATANNYNNERNNGAINNDL--KADPNDEFLKILIRSAIKCL	239
QY	241	TRPELYFVDVRSAINKTGTDEGALTRIVTTFAEIDLKIGEYORRNSIPILEKATKOT	300
Db	240	TYPEKCFEYVRLAINKGTDEBMLTRVVTTRAEVDMQRIKQYHKRNSVPLDRALANDT	299
QY	301	RGDYEKMLVALLGEDDA	317
Db	300	SGDYKKMLLALVGHEDA	316

RESULT 9  
US-10-424-599-213299

Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J

APPLICANT: Zhou yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associate

```

1 FILE REFERENCE: 38-21(53223) B
2
3 CURRENT APPLICATION NUMBER: US/10/424,599
4
5 CURRENT FILING DATE: 2003-04-28
6

```

```

; SEQ ID NO 213299
;
; LENGTH: 319
; TYPE: PRT

```

```

1  FEATURE:      unsure
2  NAME/KEY:     unsure
3  LOCATION:     (1) .. (319)
4

```

```

; FEATURE: Clone ID: PAT_MRT3847_34633C.1.pep
; OTHER INFORMATION:
US-10-424-599-213299

```

Query Match	65.5%	Score 1054.5	DB 15	Length 319
Best Local Similarity	66.8%	Pred. No. 1.6e-82		
Matches 213, Conservative	45	Mismatches 58	Indels 3	Gaps 3

```

QY      1 MATLKVSDSDYAPASPDAAEOLRTAFEGWGTNEDLIISIARSAEORKVIRQAYHETRYGED 60
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||::: 60
Db      1 MATLVPAQPLSPLEDSOEQLKAFQGGWNGNEGLISILGHNNAAQRLKLEAYSTTHGED 60

```

QY 61 LKLTIDKEIANDPERAIIILMTLPEPERDALANEATKRWTSNNOCVMEVACTRTSTQLH 120  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
Db 61 LKLTIDKEISSDPERAVLWITLDPERDAFLANEAETIKMLTSSNWVILEIASTRSSDLIK 120

QY	121	ARQAVHAR YKKSLIEDVAHHTTGDRKL VSLVTSRYEEGG - EVNNNTLAKQSAKL VHEKI	179
		121	180
Db	121	AKQAVQARFKSLIEDVAHHTTGDRKL VSLVTSRYEEGG - EVNNNTLAKQSAKL VHEKI	180

QY 180 KDKHNDVDYIRILSTRSKAOINAFNRYYDDHGEILKSE-EGDDDDKFLALRS-TI 237  
181 AKAANDDEBILIRISTRSKQTLATLNQYINEFGAINKDLKELKQVAINCLNLRAXAI 240  
Db

QY 238 QCLRLPELKEVDVRLSAINTGTBEGALTIIVTTAEIDLKVIQEYQRRSIPLEKAIT 297  
 Db 241 KCLYLPKFAKVLRLAINTGTBEGALTRVTTAEVDLQRIAEYQRRSIPLEKRAIA 300

```
Qy      298 KDRGDYEKMLVALLGEDD 316
          |||||:|:|:|:|
Db      301 SPTSGDYQSILLALVGHEH 319
```

```

RESULT 10
US-10-767-701-45089
: Sequence 45089, Application US/10767701
: Publication NO. US20040172684A1
: GENERAL INFORMATION:
: APPLICANT: Kovalic, David K.
: APPLICANT: Cao, Yongwei
: APPLICANT: Zhou, Yihua
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
: FILE REFERENCE: 38-21(53535) B
: CURRENT APPLICATION NUMBER: US/10/767, 701
: CURRENT FILING DATE: 2004-01-29
: NUMBER OF SEQ ID NOS: 63128
: SEQ ID NO 45089
: LENGTH: 314
: TYPE: PRT
: ORGANISM: Sorghum bicolor
: FEATURE:
: OTHER INFORMATION: Clone ID: SORBI-28MAY03-C5887_1_dep
: US-10-767-701-45089

```

```

Query Match 62.8%; Score 1011.5; DB 16, Length 314;
Best Local Similarity 61.0%; Pred. No. 8,2e-79;
Matches 192; Conservative 53; Mismatches 69; Indels 1; Gaps 1;

QY 1 MATLKVSDSVPAASDDAEQLRTAFEGWGTNEDLLISIIILHRSAEQRKVIROYVHETGYED 60
    ||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 1 MATLKVPATVPVADDCQGLRKAFGQWGTNELLIIYISLGHRAAQRRARIRAYVAEHGHE 60
    ||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY 61 LKLTIDKLSLNDFFAAILLMTLEPGERDALLANEATKMTSSNOYLMEVACTRTSTQLIH 120
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 61 LLRSITDEISGFEERAVILMTLDPAPERDAVLANEARKKQPNRVLVEIACRTSQAQVA 120
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 121 AROAVHAAYKKSLIEDVVAHHTTGDGRKLLVSLVTSYRYSGEDVWNMTLQAQKLVHEKIK 180
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 121 AROAVHEHFKSLIEDIAAHVTDGRKLLVPLVSTRYDGPVNRRLASEKLLHEKIKH 180
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 181 DKHYNDEDEVIRILSTRKAQINATNRYQODHGEELISLEGGDDDDKFTALLRSTIOCL 240
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 181 HKAYSDDSEIIRILTRRSKPKQLLATITNNHNDAGHINKDL-KADPKDEVLKTLRAIRRF 239
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 241 TRPELYFVDVLRSAINKGTGDEGALTRIVTTTAEIDLVKYGEEYORNSIPILEKAITKDT 300
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 240 SCSPRYFEKVAQQAAGLGTDENALTTRVITTTAEVDLKLKEAYOKRNSVPLERAVAGDT 299
    :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 301 RGDEYKMLVALLGED 315
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 300 SGDYESMLALLGQE 314
    ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 11
US-10-425-114-60631
; Sequence 60631, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128

```

```

; SEQ ID NO 60631
;
; LENGTH: 239
;
; TYPE: PRT
;
; ORGANISM: Zea mays
;
; FEATURES:
;
; OTHER INFORMATION: Clone ID: LIB3356-011-F3_FLI.ppt
;
; US-10-425-114-60631

```

Query Match	62.7%	Score 1009.5	DB 15	Length 339
Best Local Similarity	60.6%	Pred. No. 1.3e-78		
Matches 191, Conservative	53	Mismatches 70	Indels 1	Gaps 1

[illegible]

```

RESULT 12
US-10-425-115-186841
Sequence 186841, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 186841
LENGTH: 340
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MFT4577_101989C.1.pep
US-10-425-115-186841

```

[illegible]

```
Db      147 TRQAVHERFRKSLSEEDIAAHVTGDFRKLVLPLVSTYRYQDPEVNTRLAHSEAKLHEKXH 206
      QY      181 DKRYNDEVIRILSTSKQAQINATFNRYQDDHGEELKLSSEDDDDKFLALIRSTIOCL 240
      Db      207 HKAYSDEIIRILITTRSKPOLATFNHYDAFGHRINKOL-KADPDDEVILTRLAIRIRCF 265
      QY      241 TRPELVFVDVLRSAIKNGTDEGALTRIVTTTRABIDLKVIGEYORNSIPILEKAITKDT 300
      Db      266 SCDDRFEKVAQAQIAGLGTDENSLTRVITTRAEVDLKLKEAYQKRNVPLERAVAGDT 325
      QY      301 RGDYERKMLVALIGED 315
      Db      326 SGDYESMLLALIGQE 340

RESULT 13
US-10-425-114-64727
; Sequence 64727, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64727
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4605-011-H2_FLI.pep
US-10-425-114-64727

Query Match      62.7%; Score 1009.5; DB 15; Length 361;
Best Local Similarity 60.6%; Pred. No. 1.5e-78;
Matches 191; Conservative 53; Mismatches 70; Indels 1; Gaps 1;

QY      1 MATLKVSQSDVPASDDAQLRTAFEGWGTNEDLIISILAHRSABQKVIROAYHETGED 60
      Db      48 MATLKVPATVPVADDCDQLRKAFQGWGTNEALIIISLGRDAQRRAIRRAYAAVGE 107
      QY      61 LKTLKLEISNDFERAIIILMTLEPGRDALLANEATKRWTSNQVLMVEVACTSTQOLH 120
      Db      108 LIRSIIDELISGDFERAVILMTLDPARDAVLANEARKKPKGNRLVETLICTRTSQIFA 167
      QY      121 ARQAVHARYKSLSEEDVAHHTTGDPRKLVLVSLVTSYRYGDEEVNMTLAKQEAQLVHEKIK 180
      Db      168 TRQAVHERFRKSLSEEDIAAHVTGDFRKLVLPLVSTYRYQDPEVNTRLAHSEAKLHEKXH 227
      QY      181 DKRYNDEVIRILSTSKQAQINATFNRYQDDHGEELKLSSEDDDDKFLALIRSTIOCL 240
      Db      228 HKAYSDEIIRILITTRSKPOLATFNHYDAFGHRINKOL-KADPDDEVILTRLAIRIRCF 286
      QY      241 TRPELVFVDVLRSAIKNGTDEGALTRIVTTTRABIDLKVIGEYORNSIPILEKAITKDT 300
      Db      287 SCDDRFEKVAQAQIAGLGTDENSLTRVITTRAEVDLKLKEAYQKRNVPLERAVAGDT 346
      QY      301 RGDYERKMLVALIGED 315
      Db      347 SGDYESMLLALIGQE 361

RESULT 14
US-10-425-114-60513
; Sequence 60513, Application US/10425114
```

```
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60513
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3595-049-H3_FLI.pep
US-10-425-114-60513

Query Match      62.6%; Score 1008.5; DB 15; Length 341;
Best Local Similarity 60.6%; Pred. No. 1.7e-78;
Matches 191; Conservative 53; Mismatches 70; Indels 1; Gaps 1;

QY      1 MATLKVSQSDVPASDDAQLRTAFEGWGTNEDLIISILAHRSABQKVIROAYHETGED 60
      Db      28 MATLKVPATVPVADDCDQLRKAFQGWGTNEALIIISLGRDAQRRAIRRAYAAVGE 87
      QY      61 LKTLKLEISNDFERAIIILMTLEPGRDALLANEATKRWTSNQVLMVEVACTSTQOLH 120
      Db      88 LIRSIIDELISGDFERAVILMTLDPARDAVLANEARKKPKGNRLVETLICTRTSQIFA 147
      QY      121 ARQAVHARYKSLSEEDVAHHTTGDPRKLVLVSLVTSYRYGDEEVNMTLAKQEAQLVHEKIK 180
      Db      148 TRQAVHERFRKSLSEEDIAAHVTGDFRKLVLPLVSTYRYQDPEVNTRLAHSEAKLHEKXH 207
      QY      181 DKRYNDEVIRILSTSKQAQINATFNRYQDDHGEELKLSSEDDDDKFLALIRSTIOCL 240
      Db      208 HKAYSDEIIRILITTRSKPOLATFNHYDAFGHRINKOL-KADPDDEVILTRLAIRIRCF 266
      QY      241 TRPELVFVDVLRSAIKNGTDEGALTRIVTTTRABIDLKVIGEYORNSIPILEKAITKDT 300
      Db      267 SCDDRFEKVAQAQIAGLGTDENSLTRVITTRAEVDLKLKEAYQKRNVPLERAVAGDT 326
      QY      301 RGDYERKMLVALIGED 315
      Db      327 SGDYESMLLALIGQE 341

RESULT 15
US-10-425-114-59289
; Sequence 59289, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59289
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
```



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 20:28:24 ; Search time 179 Seconds

(without alignments)  
906.866 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611

Sequence: 1 MATLKVSDSVAPASDPAEQL.....KOTRGDYKXKLVALLGSDDA 317

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1611	100.0	317	2	09SYT0
2	1585	98.4	317	2	096527
3	1549.5	96.2	314	2	039001
4	1507	93.5	317	2	067EX8
5	1202.5	74.6	316	2	082090
6	1193.5	74.1	315	2	093157
7	1173.5	72.8	316	2	022341
8	1117.5	69.4	314	2	042657
9	1112.5	69.1	314	2	09SB88
10	1109.5	68.9	316	2	09ZRU7
11	1107	68.7	315	2	093158
12	1106.5	68.7	316	2	09XEN8
13	1096.5	68.1	314	2	024131
14	1096.5	68.1	314	2	09M3H3
15	1095.5	68.0	314	2	081536
16	1093.5	67.9	316	2	069DC2
17	1087.5	67.5	314	2	024132
18	1080	67.0	315	2	081535
19	1067.5	66.3	308	2	042922
20	1067	66.2	317	2	09XER2
21	1037.5	64.4	316	2	09LX07
22	1019.5	63.3	316	2	09CSV2
23	1009.5	62.7	314	2	043863
24	1004.5	62.4	312	2	067V21
25	991.5	61.5	318	2	0681H6
26	991.5	61.5	318	2	09LX08
27	986.5	61.2	314	2	06Z6K7
28	982.5	61.0	318	2	09CSV3
29	949.5	58.9	314	2	043864
30	879	54.6	314	1	ANX4_FRAAN
31	781.5	48.5	313	2	065848

32	715.5	44.4	323	2	0814Z2	0814Z2	oryza sativ
33	635.5	39.4	372	2	06L4C5	06L4C5	oryza sativ
34	602.5	37.4	321	2	09ZVJ7	09ZVJ7	arabidopsis
35	597.5	37.1	321	2	09SB45	09SB45	arabidopsis
36	579.5	36.0	257	2	094CK4	094CK4	arabidopsis
37	523	32.5	316	2	09C9X3	09C9X3	arabidopsis
38	522	32.4	316	2	09CSV4	09CSV4	arabidopsis
39	512	31.8	315	2	06H450	06H450	oryza sativ
40	511.5	31.8	333	2	09ZRS3	09ZRS3	medicago sa
41	505	31.3	316	2	06FHH6	06FHH6	homo sapien
42	502	31.2	483	2	07T391	07T391	brachydanio
43	502	31.2	526	2	0804G4	0804G4	brachydanio
44	501.5	31.1	303	2	08H5B8	08H5B8	oryza sativ
45	501	31.1	315	1	ANX4_HUMAN	P27216	homo sapien

#### ALIGNMENTS

RESULT 1	ID	Q9SYT0	PRELIMINARY;	PRT;	317 AA.
AC	Q9SYT0				
DT	01-MAY-2000	(TrEMBLrel. 13, Created)			
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)			
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)			
DE	Ca2+-dependent membrane-binding protein annexin (F14D7.2 protein).				
GN	Name=Annx1; Synonyms=At1g35720, F14D7.2;				
OS	Arabidopsis thaliana (Mouse-ear cress);				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	euroside II; Brassicales; Brassicaceae; Arabidopsis.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Clark G.B., Roux S.J.;				
RT	"Isolation and Characterization of Two Different Arabidopsis Annexin				
RT	cDNAs (Accession Nos. AF083913 and AF083914) (PGR 99-065).";				
RL	Plant Physiol. 120:340-340(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Nuygen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,				
RA	Palm C.U., Bowser L., Jones T., Bann J., Carninci P., Chen H.,				
RA	Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,				
RA	Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,				
RA	Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,				
RA	Ecker J., Theologis A., Davis R.W.;				
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=2088475; PubMed=12093376;				
RA	Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,				
RA	Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;				
RT	"Full-length messenger RNA sequences greatly improve genome				
RT	annotation.";				
RL	Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RA	Broyer V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,				
RA	Feldmann K.;				
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RA	Liu S.X., Chan A.C., Sakano H., Yu G., Lee J.M., Lenz C., Pham P.,				
RA	Toriumi M., Chin C., Chion J., Choi E., Chung M., Gonzalez A.,				
RA	Hwang B., Liu A., Vaynsberg M., Alcafi H., Brooks S., Buehler B.,				
RA	Chao Q., Conn L., Conway A.B., Hansen N.F., Johnson-Hopson C.,				
RA	Khan S., Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shim P.,				
RA	Southwick A., Davis R.W., Ecker J.R., Federpfeil N.A., Theologis A.,				
RA	Yu G.;				
RN	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.				
RN	[6]				

RP SEQUENCE FROM N.A.  
RA Theologis;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
RN  
RP SEQUENCE FROM N.A.  
RA Yamada K., Dale J.M., Hsuan V.W., Onodera C.S., Quach H.L.,  
RA Tortum M., Wong C., Yu H.C., Yuan S., Chen H., Cheuk R.,  
RA Jones T., Kim C.J., Nguyen M., Palm C.J., Shinn P., Southwick A.,  
RA Tripp M.G., Wu T., Davis R.W., Becker J.R., Theologis A.,  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
RN  
RP SEQUENCE FROM N.A.  
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,  
RA Carinci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,  
RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ichida J., Kamiya A.,  
RA Kawai J., Kim C.J., Natusaka M., Onodera C.S., Quach H.L., Sakurai T.,  
RA Satou M., Seki M., Shinn P., Tang C.C., Tortum M., Wong C., Yu H.C.,  
RA Yamada K., Yu G., Yuan S., Shinzaki K., Becker J., Theologis A.,  
RA Davis R.W.,  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to the annexin family.  
CC -1- SIMILARITY: Contains 4 annexin repeats.  
DR EMBL; AF083913; AAD34236.1; -.  
DR EMBL; AY072347; AAL61994.1; -.  
DR EMBL; AY086570; AAM63633.1; -.  
DR EMBL; AC021198; AAF79882.1; -.  
DR EMBL; AF323435; AAG48798.1; -.  
DR EMBL; BT003359; AAO29977.1; -.  
DR FIRM; C86479; C86479.  
DR HSSP; P93157; INO0.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.  
DR InterPro; IPR001464; Annexin.  
DR InterPro; IPR009118; Annexin\_plant.  
DR Pfam; PF00191; Annexin; 4.  
DR PRINTS; PR00196; ANNEXINPLANT.  
DR PRODOM; PD000143; Annexin; 4.  
DR SMART; SM00335; ANX; 4.  
DR PROSITE; PS00223; ANNEXIN; 1.  
KW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;  
KW Repeat.  
SQ SEQUENCE 317 AA; 36203 MW; 92516D630325005F CRC64;  
Query Match 100.0%; Score 1611; DB 2; Length 317;  
Best Local Similarity 100.0%; Pred. No. 4, 2e-94;  
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MATLKVSQSVPAAPSDAEQRLTAPEGWTGNETDIIISLAHSAEQRKVIRQAVHETGSD 60  
DB 1 MATLKVSQSVPAAPSDAEQRLTAPEGWTGNETDIIISLAHSAEQRKVIRQAVHETGSD 60  
QY 61 LKLTLDKELSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMEVACTRTSTQQLH 120  
DB 61 LKLTLDKELSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMEVACTRTSTQQLH 120  
QY 121 AROAYHARYKKSLEEDVAHHTTGDPRKLVSVTSYRYEGDEVNMTLAKQEAQVHEKIK 180  
DB 121 AROAYHARYKKSLEEDVAHHTTGDPRKLVSVTSYRYEGDEVNMTLAKQEAQVHEKIK 180  
QY 181 DKHYNDEVDVIRILSTRSKAQINATFNRYQDDHGEILKSLBEGDDDDKFLALRSTIOCL 240  
DB 181 DKHYNDEVDVIRILSTRSKAQINATFNRYQDDHGEILKSLBEGDDDDKFLALRSTIOCL 240  
QY 241 TRPELYFVDVLRSAINKTGTDEGALTRIYTTTAEIDLKVIGGEYQRNSIPLEKAITKDT 300  
DB 241 TRPELYFVDVLRSAINKTGTDEGALTRIYTTTAEIDLKVIGGEYQRNSIPLEKAITKDT 300  
QY 301 RGDYEKMLVALGEGDDA 317  
DB 301 RGDYEKMLVALGEGDDA 317

RESULT 2  
ID 096527 PRELIMINARY; PRT; 317 AA.  
AC 096527;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Arabidopsis thaliana annexin-like protein.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC euroids II; Brassicales; Brassicaceae; Arabidopsie.  
OX NCBI\_Taxid=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=whole seedling;  
RA MEDLINE=97008169; PubMed=8855345; DOI=10.1073/pnas.93.20.11268;  
RX Gidrol X., Sabelli P.A., Fern Y.S., Kush A.K.,  
RT "Annexin-like protein from Arabidopsis thaliana rescues delta oxyR  
mutant of Escherichia coli from H2O2 stress."  
RL Proc. Natl. Acad. Sci. U.S.A. 93:11268-11273(1996).  
CC -1- SIMILARITY: Belongs to the annexin family.  
CC -1- SIMILARITY: Contains 4 annexin repeats.  
DR EMBL; U28415; AAC49472.1; -.  
DR HSSP; P93157; INO0.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.  
DR InterPro; IPR001464; Annexin.  
DR Pfam; PF00191; Annexin; 4.  
DR PRINTS; PR00196; ANNEXIN.  
DR PRODOM; PD000143; Annexin; 4.  
DR SMART; SM00335; ANX; 4.  
DR PROSITE; PS00223; ANNEXIN; 1.  
KW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;  
KW Repeat.  
SQ SEQUENCE 317 AA; 36172 MW; 375E9FA536E0D384 CRC64;  
Query Match 98.4%; Score 1585; DB 2; Length 317;  
Best Local Similarity 98.4%; Pred. No. 1, 8e-92;  
Matches 312; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MATLKVSQSVPAAPSDAEQRLTAPEGWTGNETDIIISLAHSAEQRKVIRQAVHETGSD 60  
DB 1 MATLKVSQSVPAAPSDAEQRLTAPEGWTGNETDIIISLAHSAEQRKVIRQAVHETGSD 60  
QY 61 LKLTLDKELSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMEVACTRTSTQQLH 120  
DB 61 LKLTLDKELSNDFERRAILMTLEPGERDALLANEATKRTSSNOVLMEVACTRTSTQQLH 120  
QY 121 AROAYHARYKKSLEEDVAHHTTGDPRKLVSVTSYRYEGDEVNMTLAKQEAQVHEKIK 180  
DB 121 AROAYHARYKKSLEEDVAHHTTGDPRKLVSVTSYRYEGDEVNMTLAKQEAQVHEKIK 180  
QY 181 DKHYNDEVDVIRILSTRSKAQINATFNRYQDDHGEILKSLBEGDDDDKFLALRSTIOCL 240  
DB 181 DKHYNDEVDVIRILSTRSKAQINATFNRYQDDHGEILKSLBEGDDDDKFLALRSTIOCL 240  
QY 241 TRPELYFVDVLRSAINKTGTDEGALTRIYTTTAEIDLKVIGGEYQRNSIPLEKAITKDT 300  
DB 241 TRPELYFVDVLRSAINKTGTDEGALTRIYTTTAEIDLKVIGGEYQRNSIPLEKAITKDT 300  
QY 301 RGDYEKMLVALGEGDDA 317  
DB 301 RGDYEKMLVALGEGDDA 317  
RESULT 3  
ID 039001 PRELIMINARY; PRT; 314 AA.  
AC 039001;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

```

DE Annexin (fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Schantz R., Schantz M.L., Houline G.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the annexin family.
CC -1- SIMILARITY: Contains 4 annexin repeats.
DR EMBL; X99224; CAA67608.1; -.
DR HSSP; P93157; IN00.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR009118; Annexin_plant.
DR Pfam; PF00191; Annexin_4.
DR PRINTS; PR00196; ANNEXINPLANT.
DR PRODOM; PD000143; Annexin; 4.
DR SMART; SM00335; ANX; 4.
DR - Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;
KM Repeat.
FT NON TER 1 1
SQ SEQUENCE 314 AA; 35779 MW; D617ADCAB8E4F91 CRC64;

Query Match 96.2%; Score 1549.5; DB 2; Length 314;
Best Local Similarity 97.5%; Pred. No. 3.2e-90;
Matches 307; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 3 TLAKVSDVPAPSDAEQLRTAFEGMGNTMEDLIISILAHSAEQKRVIRQAVHETGYGDL 62
DB 1 TLAKVSDVPAPSDAEQLRTAFEGMGNTMEDLIISILAHSAEQKRVIRQAVHETGYGDL 60
QY 63 KTLDKESLNDFERRAILMTLEPGERDALIANEATKRTSSNOVMEVACTRTSTQILHAR 122
DB 61 KTLDKESLNDFERRAILMTLEPGERDALIANEATKRTSSNOVMEVACTRTSTQILHAR 120
QY 123 QAVHARYKKSLEEDVAHHTTGDFFKLLVSLVTSRYRGDEVNMTLAKQEAQLVHEKIKDK 182
DB 121 QAVHARYKKSLEEDVAHHTTGDFFKLLVSLVTSRYRGDEVNMTLAKQEAQLVHEKIKDK 180
QY 183 HYNDEDEVIRILSTSKAQINATFNRYODDNGEELKSLBEGDDDKFLALRSTIOCLTR 242
DB 181 HYNDEDEVIRILSTSKAQINATFNRYODDNGEELKSLBEGDDDKFLALRSTIOCLTR 240
QY 243 PELYFVVLRSAINKTGTDEGALTRIVTTRAEIDLKVIIGEYORRNSIPILEKAITKOTRG 302
DB 241 PELYFVVLRSAINKTGTDEGALTRIVTTRAEIDLKVIIGEYORRNSIPILEKAITKOTRG 299
QY 303 DYERKMLVALGEGDDA 317
DB 300 DYERKMLVALGEGDDA 314

```

```

RA Jami S.K., Kirti P.B.;
RT "Annexin cDNA from the floral bud of Brassica juncea.";
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the annexin family.
CC -1- SIMILARITY: Contains 4 annexin repeats.
DR EMBL; AY563355; AAR10457.1; -.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR009118; Annexin_plant.
DR Pfam; PF00191; Annexin_4.
DR PRINTS; PR00196; ANNEXIN.
DR PRODOM; PD000143; Annexin; 4.
DR SMART; SM00335; ANX; 4.
DR PROSITE; PS00223; ANNEXIN.1.
KM Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;
KM Repeat.
SQ SEQUENCE 317 AA; 36220 MW; 1A4F9211AAF39A2A CRC64;

Query Match 93.5%; Score 1507; DB 2; Length 317;
Best Local Similarity 91.2%; Pred. No. 1.6e-87;
Matches 289; Conservative 22; Mismatches 6; Indels 0; Gaps 0;

QY 1 MATLKVSQSDVPAPSDAEQLRTAFEGMGNTMEDLIISILAHSAEQKRVIRQAVHETGYGDL 60
DB 1 MATLKVSQSDVPAPSDAEQLRTAFEGMGNTMEDLIISILAHSAEQKRVIRQAVHETGYGDL 60
QY 61 LKTLDKESLNDFERRAILMTLEPGERDALIANEATKRTSSNOVMEVACTRTSTQILHAR 120
DB 61 LKTLDKESLNDFERRAILMTLEPGERDALIANEATKRTSSNOVMEVACTRTSTQILHAR 120
QY 121 ARQAVHARYKKSLEEDVAHHTTGDFFKLLVSLVTSRYRGDEVNMTLAKQEAQLVHEKIKDK 180
DB 121 ARQAVHARYKKSLEEDVAHHTTGDFFKLLVSLVTSRYRGDEVNMTLAKQEAQLVHEKIKDK 180
QY 181 DKNYNDEDEVIRILSTSKAQINATFNRYODDNGEELKSLBEGDDDKFLALRSTIOCLTR 240
DB 181 DKNYNDEDEVIRILSTSKAQINATFNRYODDNGEELKSLBEGDDDKFLALRSTIOCLTR 240
QY 241 TRPELYFVVLRSAINKTGTDEGALTRIVTTRAEIDLKVIIGEYORRNSIPILEKAITKOT 300
DB 241 TRPELYFVVLRSAINKTGTDEGALTRIVTTRAEIDLKVIIGEYORRNSIPILEKAITKOT 300
QY 301 RGDYERKMLVALGEGDDA 317
DB 301 RGDYERKMLVALGEGDDA 317

```

```

RESULT 4
Q6TEX8 PRELIMINARY; PRT; 317 AA.
AC 06TEX8;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Annexin.
OS Brassica juncea (leaf mustard) (Indian mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3707;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=floral bud;

```

```

RESULT 5
O82090 PRELIMINARY; PRT; 316 AA.
AC 082090;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Fiber annexin.
OS Gossypium hirsutum (upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Malvales; Malvaceae; Malvaceae; Gossypium.
OX NCBI_TaxID=3635;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Texas Marker1;
RA Shin H., Brown R.M.Jr.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the annexin family.
CC -1- SIMILARITY: Contains 4 annexin repeats.
DR EMBL; U89609; AAC33305.1; -.
DR FJR; J31428; J31428.
DR HSSP; P93157; IN00.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.
DR InterPro; IPR001464; Annexin.

```

DR	ProDom: PD000143; Annexin, 4.
DR	SMART; SM00335; ANX, 4.
DR	PROSITE; PS00223; ANNEXIN, 1.
KW	Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;
KW	Repeat.
FT	NON TER
SO	SEQUENCE 1 1 35823 MW; 653113C77AAAB980 CRC64;
Query Match	74.1%; Score 1193.5; DB 2; Length 315;
Best Local Similarity	73.4%; Pred. No. 1e-67;
Matches 229;	Conservative 37; Mismatches 45; Indels 1; Gaps 1;
QY	2 ATKLVSDVSPAPSDPAEQRLRTAFEGWCTNEDLIISILHRSAEQKVIQVAYHETYGEDL 61
DB	1 ATLTVPTTVSVSDSCQLKAKFSQMGNTGILLIDIGHMAQRNLIRKTYETIYGEDL 60
QY	62 LKTLIDKELSNDFERRAILLMTLPEGRDALLANETKRWISSNOVIMEVACTRTSTOLLHA 121
DB	61 LKALDKELSNDFERRVLLMLDPAERDALLANETKRWISSNOVIMEIACRASNOLLHA 120
QY	122 RQAHAAVKKSLSEDDVAHHTGDPFRKLVSLNYSYRREGSEVMNTLAKQAKIVHEKTD 181
DB	121 RQAHAAVKKSLSEDDVAHHTGDPFRKLLPLVSSYRREGSEVMNTLAKTAKLHEKTSN 180
QY	182 KHVNDVEDIRILSTRKQAINATFNRYQDDHGEELKSLSEGGDDDDKFTLALIRSTIQLCT 241
DB	181 KASDDDDVIVLAVTRSAQGINATLNHKNYGNINDKL-KADPKDFLALLASTYKCV 239
QY	242 RPELYFVDVRSAINKTGTDEGALTRLVTTRAEIDLKVGEEVQRRNSITLEQLITKOTR 301
DB	240 YPEKYCFEKVRLAIRNGRTDEGALTRVCTRAEVDLKVIADEYQRRNSVPLTRAIYVDTH 299
QY	302 GDYERKMLVALLG 313
DB	300 GDYERKMLVALLG 311
RESULT 7	
ID	022341 PRELIMINARY; PRT; 316 AA.
AC	022341
DT	01-JAN-1998 (TREMBLrel_05, Created)
DT	01-JAN-1998 (TREMBLrel_05, Last sequence update)
DT	01-MAR-2004 (TREMBLrel_26, Last annotation update)
DE	Annexin.
GN	Name=Annx1;
OS	Lavatera thuringiaca.
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC	eurosid II; Malvales; Malvaceae; Malvaceae; Lavatera.
OX	NCBI_TaxID=61660;
RN	[1]
RN	SEQUENCE FROM N.A.
RC	TISSUE=Leaves without petioles;
RA	Vazquez-Tello A., Uozumi T.,
RT	"Cloning and characterization of a Lavatera thuringiaca cDNA encoding
RT	an annexin whose expression is stimulated by low temperature.";
RL	Plant Physiol. 0:0-0(1997).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Leaves without petioles;
RX	MEDLINE=20255875; PubMed=10795312;
RA	Breton G., Vazquez-Tello A., Danyluk J., Sarhan F.;
RT	"Two novel intrinsic annexins accumulate in wheat membranes in
RT	response to low temperature.";
CC	Plant Cell Physiol. 41:177-184(2000).
CC	-1- SIMILARITY: Belongs to the annexin family.
CC	-1- SIMILARITY: Contains 4 annexin repeats.
DR	EMBL; AF006197; AAB71830.1; -.
DR	HSEF; P93157; IM00.
DR	GO; GO:0005509; F:calcium ion binding; IEA.
DR	GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.
DR	InterPro; IP0001464; Annexin.

Query Match	Similarity	Score	DB 2	Length	316
Beat Local	71.0%	Pred. No. 1.9e-66;			
Matches	225; Conservative	43; Mismatches	48; Indels	1; Gaps	14
DR	InterPro: IPR009118; Annexin_plant.				
DR	Pfam: PF00191; Annexin; 4.				
DR	PRINTS; PR00196; ANNEXIN.				
DR	PRINTS; PR01814; ANNEXINPLANT.				
DR	Prodom; PD000143; Annexin; 4.				
DR	SMART; SM00335; ANX; 4.				
DR	PROSITE; PS00223; ANNEXIN; 1.				
KM	Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding; Repeat.				
SO	SEQUENCE	316 AA; 36083 MW; EAB906686A351AA CRC64;			
Qy	1	MATLVKVSVPAPSDAEQLRTAFEGMGTNEDLIISILAHRSAEQRVIPOAHETYGED	60		
Db	1	MATLVPSPTLVSVSDECEQLRAFGMGKTNEDLIINIGHRRADRNIRKRAYTTHGED	60		
Qy	61	LLKTKDELKSNDFERRAILLMTLEPERDALLANETKWTSSNQVAMEVACTRISTOLLH	120		
Db	61	LLKALDELKSNDFELVLLMTLDPEDDALLANETKWTSSNQVIMEIACRSSDQLLR	120		
Qy	121	ARQAVHAKYKSLSEEDVAHHTTGGDRKLLVSLVTSYRREGSDRVNMTLAKQKALVHEKIK	180		
Db	121	ARQAVHVKYKSLSEEDVAHHTTGGDRKLLPLVSSYRYEGEVMNTLAKTEKALHEKIS	180		
Qy	181	DKHAVDEIVIRILSTRSKAQINATFNRYODHGEIILKSLSEGGDDDKFLALRSTIOCL	240		
Db	181	NKASDDDIVIVLARSKSQINERLNHYKNEATDINDKL-KADPKDEFLLALSTVACL	239		
Qy	241	TRPELVFVDVILRSAINKTGTDEGALTIRIVTTAEIDLKVIAGEYGRNRSIPLEKATYOT	300		
Db	240	VYPEKVFPEKVLRLAINKGTDEGALTIRVVSFRAEVDLKIADVEYGRNRSVPLTRALVOT	299		
Qy	301	RGDYERKMLVALGDEDA	317		
Db	300	NGDYERKLLVLAAGEVEA	316		
RESULT 8					
Q42657	PRELIMINARY;	PRT;	314 AA.		
AC	Q42657;				
DT	01-NOV-1996 (TREMBLrel. 01, Created)				
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)				
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)				
OS	Annexin.				
OS	Capiscum annum (Bell pepper).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;				
OC	lamids; Solanales; Solanaceae; Capsicum.				
OX	NCBI_TaxId=4072;				
AP	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=fruit;				
RX	MEDLINE=96159603; PubMed=8925897; DOI=10.1016/0014-5793(96)00252-9;				
RA	Proust J., Houle G., Schantz M.L., Schantz R.;				
RT	"Characterization and gene expression of an annexin during fruit				
RT	development in Capsicum annum.";				
CC	FEBS Lett. 383:208-212(1996).				
CC	-1- SIMILARITY: Belongs to the annexin family.				
DR	-1- SIMILARITY: Contains 4 annexin repeats.				
DR	EMBL; X93308; CA66710.1; .				
DR	PIR; S66274; S66274.				
DR	PDB; 1DK5; X-ray; A/B=1-314.				
DR	GO; GO:0005509; F:calcium ion binding; IEA.				
DR	GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.				
DR	InterPro; IPR001164; Annexin.				
DR	InterPro; IPR009118; Annexin_plant.				
DR	Pfam; PF00191; Annexin; 4.				
DR	PRINTS; PR00196; ANNEXIN.				
DR	PRINTS; PR01814; ANNEXINPLANT.				

DR	PRODUM; PP0000143; Annexin; 4.
DR	SMART; SM00335; ANX; 4.
DR	PROSITE; PS00223; ANNEXIN; 1.
KW	Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;
KW	Repeat.
SQ	SEQUENCE 314 AA; 35850 MW; 891A8286C4A029DB CRC64;
Query Match	69.4%; Score 1117.5; DB 2; Length 314;
Best Local Similarity	67.1%; Pred. No. 6.6e-63;
Matches 212; Conservative 58; Mismatches 43; Indels 3; Gaps 1;	
OY	1 MATLKVSDDSPAPSDDADEQLRTAFEGSGTTEEDLIISILARSAPQRVIRQAHYETGED 60
DB	1 MASLTVAHPVSAADEQOLRSARFGKGTVEKKIISIILARRTAORLIKQTVAETGECD 60
OY	61 LKLTKDELSEIDPFRALITLMTLBEGEEDALLANEATGRWTSSNOVLMEVACTRSTOLLH 120
DB	61 LLKLELDRLTHDFEKLVLTWLTPDSEDAHLAKAEATGRWKSNFVLVELLACTRSPKEVLV 120
OY	121 AROAYHARYKKSLEEDVAHHHTGDFRKLLVSLVTSYREBGEDEVNMLTAKOBAKVHEKIK 180
DB	121 AREAVHARYKKSLEEDVAHYHTGDHGRKLVLPLVSYRYGGAEDVDLRKAESKIHEKIS 180
OY	181 DKAHNDEDVRIILSTRSKAQIINATFNRYODDHGEELIKSLEEGDDDKFLALRSTIQCL 240
DB	181 DKAVSDDEVIRILATRASKAQIINTLNMYKHDEHGEDIKOLEDGDE--PALLRATIKGL 237
OY	241 TRPLEYEVADVLSAINKTGTDEGALTIRIVTTPAEIDLKVIAGEFYORNSIPLEKAIPTOT 300
DB	238 VYEPHYVEVRDANRGTFEBDLTRVIAITRAEVDLKITIADERYKDSDIPIGRALKANDT 297
OY	301 RGDYEKMVLVALLGEDD 316
DB	298 RGDYESMVLVALLGEE 313
RESULT 9	
ID	O9SB88 PRELIMITARY; PRT; 314 AA.
AC	O9SB88:
DT	01-MAY-2000 (TREMBLrel_13, Created)
DT	01-MAY-2000 (TREMBLrel_13, Last sequence update)
DT	01-MAR-2004 (TREMBLrel_26, Last annotation update)
DE	Annexin cap32.
GN	Name=an.1;
OS	Capsicum annuum (Bell pepper).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; asterids;
CC	lamiales; Solanales; Solanaceae; Capsicum.
OX	NCBI_TaxId=4072;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Proust J., Schantz R.;
Submitted BL	Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
-1-	SIMILARITY: Belongs to the annexin family.
-1-	SIMILARITY: Contains 4 annexin repeats.
DR	HSSP; Q42657; IDS.
GO	GO: GO:0005509; F:calcium ion binding; IEA.
GO	GO: GO:0005544; F:calcium-dependent phospholipid binding; IEA.
InterPro	InterPro; IPR001464; Annexin.
InterPro	InterPro; IPR009118; Annexin_plant.
Pfam	Pfam; PF00191; Annexin_4.
PRINTS	PRINTS; PRO0196; ANNEXIN.
PRINTS	PRINTS; PRO1814; ANNEXINPLANT.
ProDom	ProDom; PD000143; Annexin; 4.
SMART	SMART; SM00335; ANX; 4.
PROSITE	PS00223; ANNEXIN; 1.
KW	Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;
KW	Repeat.
SQ	SEQUENCE 314 AA; 35858 MW; 58CG9B8FA0A4CDDC CRC64;
Query Match	69.1%; Score 1112.5; DB 2; Length 314;



Db 61 LKCEKELTSDFERAVLFTLDPAPEROAHLANEATKXKFTSSNWTLMELIACRSHELLANV 120  
Qy 122 RQAVHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYRGEDEVNNTLAKQOAKLVHEKIKD 181  
Db 121 KKVHARYKKSLEEDVAHHTTGEYRKLLVPLVSAFRYEGDEVNNTLAKSEAKILHDKISD 180  
Qy 182 KHVNDDEVIRILSTSRKAQINATFNRYQDDHGEELKSLSEGGDDDDKFLALRSTIOCLT 241  
Db 181 KHVNDDEVIRILSTSRKAQINATFNRYQDDHGEELKSLSEGGDDDDKFLALRSTIOCLT 239  
Qy 242 RPELVFVDVLRSAINKTGDEGALTRIVTTFAEIDLKVGIEYORRNSIPLKATTKOTR 301  
Db 240 TPEGVFVKVLRQAINKLGSEWMLTRVVTTRAEVDWRIKAYORRNSIPLKQAIADTS 299  
Qy 302 GDYERKMLVALLGEDDA 317  
Db 300 GDYERKMLVALLGEDDA 315  
RESULT 12  
ID Q24131 PRELIMINARY; PRT; 316 AA.  
AC Q24131  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Vacuole-associated annexin VcAB42.  
GN Name=Anx3;  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Nicotiana.  
OX NCBI\_TaxID=4097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bright Yellow-2;  
RA Seale D.F., Randall S.K.;  
RT "Sequence Analysis of a Vacuole-Associated Annexin from Tobacco  
Plant Physiol. 119:1147-1147(1999)."  
RL Plant Physiol. 119:1147-1147(1999).  
CC -1- SIMILARITY: Belongs to the annexin family.  
CC -1- SIMILARITY: Contains 4 annexin repeats.  
DR EMBL; AF113545; AAD24540.1; -.  
DR HSSP; P93157; IM00.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.  
DR InterPro; IPR001464; Annexin.  
DR InterPro; IPR009118; Annexin\_plant.  
DR Pfam; PF00191; Annexin\_4.  
DR PRINTS; PR00196; ANNEXIN.  
DR PRINTS; PR01814; ANNEXINPLANT.  
DR ProDom; PD000143; Annexin\_4.  
DR SMART; SM00335; ANX; 4.  
DR PROSITE; PS00223; ANNEXIN; 1.  
KW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;  
KW Repeat.  
SQ SEQUENCE 316 AA; 35937 MW; 99E7973FC04C5FBE CRC64;  
Query Match 68.7%; Score 1106.5; DB 2; Length 316;  
Best Local Similarity 67.5%; Pred. No. 3,3e-67;  
Matches 214; Conservative 43; Mismatches 59; Indels 1; Gaps 1;  
Qy 1 MATLKVSDVPAAPSDAEOALRTAEFGWGTNEDLIIISILHARSAEORKVIRQAVHETYGED 60  
Db 1 MASLKVTSVPEPEADEOQLKAFAGWGTNEALIIQLAHRNAQORQLIRETYAAAVGED 60  
Qy 61 LKLTLDKELSDNFERAILMTLLEPERGDRALLANEATKRWTSNOVLMVEVACTRSTOLH 120  
Db 61 LKLTLDKELSDNFERAILMTLLEPERGDRALLANEATKRWTSNOVLMVEVACTRSTOLH 120  
Qy 61 LKLTLDKELSDNFERAILMTLLEPERGDRALLANEATKRWTSNOVLMVEVACTRSTOLH 120  
Db 61 LKLTLDKELSDNFERAILMTLLEPERGDRALLANEATKRWTSNOVLMVEVACTRSTOLH 120  
Qy 121 ARQAVHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYRGEDEVNNTLAKQOAKLVHEKIK 180  
Db 121 ARQAVHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYRGEDEVNNTLAKQOAKLVHEKIK 180

Qy 181 DKAVNDEVIRILSTSRKAQINATFNRYQDDHGEELKSLSEGGDDDDKFLALRSTIOCL 240  
Db 181 DKAVNDEVIRILSTSRKAQINATFNRYQDDHGEELKSLSEGGDDDDKFLALRSTIOCL 239  
Qy 241 TPELVFVDVLRSAINKTGDEGALTRIVTTFAEIDLKVGIEYORRNSIPLKATTKOTR 300  
Db 240 KTEPKHEKVLRLAINKTGDEGALTRIVTTFAEVDWRIKEEYHKNNSVPLRLAAGDT 299  
Qy 301 RGDYERKMLVALLGEDDA 317  
Db 300 RGDYERKMLVALLGEDDA 316  
RESULT 13  
ID Q24131 PRELIMINARY; PRT; 314 AA.  
AC Q24131  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE P32.1 annexin.  
GN Name=X511; Synonym=an.1;  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Nicotiana.  
OX NCBI\_TaxID=4097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bright Yellow 2;  
RA MEDLINE=918786; PubMed=10080701; DOI=10.1023/A:1006199814795;  
RA Prouet J., Houine G., Schantz M.L., Shen W.H., Schantz R.;  
RT "Regulation of biosynthesis and cellular localization of Sp32 annexins  
in tobacco BY2 cells."  
RL Plant Mol. Biol. 39:361-372(1999).  
CC -1- SIMILARITY: Belongs to the annexin family.  
CC -1- SIMILARITY: Contains 4 annexin repeats.  
DR EMBL; Y14972; CA475213.1; -.  
DR EMBL; Y17502; CA476769.1; -.  
DR HSSP; Q42657; IDK5.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0005544; F:calcium-dependent phospholipid binding; IEA.  
DR InterPro; IPR001464; Annexin.  
DR InterPro; IPR009118; Annexin\_plant.  
DR Pfam; PF00191; Annexin\_4.  
DR PRINTS; PR00196; ANNEXIN.  
DR PRINTS; PR01814; ANNEXINPLANT.  
DR ProDom; PD000143; Annexin\_4.  
DR SMART; SM00335; ANX; 4.  
DR PROSITE; PS00223; ANNEXIN; 1.  
KW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding;  
KW Repeat.  
SQ SEQUENCE 314 AA; 35855 MW; 9DB6F1AB3392D587 CRC64;  
Query Match 68.1%; Score 1096.5; DB 2; Length 314;  
Best Local Similarity 65.8%; Pred. No. 1.4e-61;  
Matches 208; Conservative 59; Mismatches 46; Indels 3; Gaps 1;  
Qy 1 MATLKVSDVPAAPSDAEOALRTAEFGWGTNEDLIIISILHARSAEORKVIRQAVHETYGED 60  
Db 1 MASLKVTSVPEPEADEOQLKAFAGWGTNEALIIQLAHRNAQORQLIRETYAAAVGED 60  
Qy 61 LKLTLDKELSDNFERAILMTLLEPERGDRALLANEATKRWTSNOVLMVEVACTRSTOLH 120  
Db 61 LKLTLDKELSDNFERAILMTLLEPERGDRALLANEATKRWTSNOVLMVEVACTRSTOLH 120  
Qy 61 LKLTLDKELSDNFERAILMTLLEPERGDRALLANEATKRWTSNOVLMVEVACTRSTOLH 120  
Db 61 LKLTLDKELSDNFERAILMTLLEPERGDRALLANEATKRWTSNOVLMVEVACTRSTOLH 120  
Qy 121 ARQAVHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYRGEDEVNNTLAKQOAKLVHEKIK 180  
Db 121 ARQAVHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYRGEDEVNNTLAKQOAKLVHEKIK 180  
Qy 181 DKAVNDEVIRILSTSRKAQINATFNRYQDDHGEELKSLSEGGDDDDKFLALRSTIOCL 240  
Db 181 DKAVNDEVIRILSTSRKAQINATFNRYQDDHGEELKSLSEGGDDDDKFLALRSTIOCL 240



DB 181 DKAYSDEVIIRILATRSKAQINATLNHYKDEVEDILKOLEEGDE---FVGLRATIKGL 237

QY 241 TRPELYFVDVLRSAIKKTGDEGALTRIVTTTAEIDLVKVGEEYORNSIPILEKATTKDT 300

DB 238 VYBEHFVEVLRDAINRGTEDEHLLTRVATRAEVDVKIADRYOKRDSIPIGRALAKDT 297

QY 301 RGDYERKMLVALLGEDD 316

DB 298 RGDYESMLLALIGQEE 313

RESULT 14

Q9M3H3 PRELIMINARY; PRT; 314 AA.

AC Q9M3H3; ID

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Annexin p34.

GN Name=an34;

OS Solanum tuberosum (Potato).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;

OC Lamiales; Solanales; Solanaceae; Solanum.

OX NCBI\_TaxID=4113;

RN [1]

RP SEQUENCE FROM N.A.

RC Kandia R., Feusner I.,

RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the annexin family.

CC -1- SIMILARITY: Contains 4 annexin repeats.

DR EMBL; AJ401032; CAB92956.1; -.

DR HSSP; Q42657; 1DK5.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR GO; GO:0005444; F:calcium-dependent phospholipid binding; IEA.

DR InterPro; IPR001464; Annexin.

DR InterPro; IPR009118; Annexin\_plant.

DR Pfam; PF00191; Annexin\_4.

DR PRINTS; PR00196; ANNEXIN.

DR PRINTS; PR01814; ANNEXINPLANT.

DR ProDom; PD000143; Annexin\_4.

DR SMART; SM00335; ANX; 4.

DR PROSITE; PS00223; ANNEXIN; 1.

DR KMW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding; Repeat.

QY \* SEQUENCE 314 AA; 35845 MW; 9E63375C356F8C36 CRC64;

Query Match 68.1%; Score 1096.5; DB 2; Length 314;

Best Local Similarity 66.5%; Pred. No. 1.4e-61;

Matches 210; Conservative 56; Mismatches 47; Indels 3; Gaps 1;

QY 1 MATLKVSQSDVPAASDADQRLTAFFEGMGNTEDLIISILAHRSABQRKVIROYAHHETGSD 60

DB 1 MASLTVPAAEVPVSAEDCEQLRSAPKGGWINEKLIISILAHRNAQRKILROYAETFGSD 60

QY 61 LKLTLDKELSNDEPERAILLWTLPEGERDALLANEAATKWTSSNOVMEVACTRTSTOLH 120

DB 61 LKELDRRLTHDFEKLVLWTLDPSESDAVLAKEATKWTKSNFVLVEIACRTSRKELVL 120

QY 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYRYEGDEVNMTLAKOEAQVHEKIK 180

DB 121 AREAYHARNRKKSLEEDVAHHTTGDPRKLLVPLVSSYRYGDEVDLRLAKAESKVLHEKIS 180

QY 181 DKRYNDEDEVIRILSTRSKAQINATFNRYODDGEELIKSLSEGGDDDKFLALIRSTIOCL 240

DB 181 DKRYSDDEVIRILATRSKAQINATLNHYKDEVEDILKOLE---DDEDFVALLRATIKGL 237

QY 241 TRPELYFVDVLRSAIKKTGDEGALTRIVTTTAEIDLVKVGEEYORNSIPILEKATTKDT 300

DB 238 VYBEHFVEVLRDAINRGTEDEHLLTRVATRAEVDVKIADRYOKRDSIPIGRALAKDT 297

QY 301 RGDYERKMLVALLGEDD 316

DB 298 RGDYENMLVALIGQEE 313

RESULT 15

ID 081536 PRELIMINARY; PRT; 314 AA.

AC 081536; ID

DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Annexin p34.

GN Name=AN34;

OS Lycopersicon esculentum (Tomato).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;

OC Lamiales; Solanales; Solanaceae; Solanum.

OX NCBI\_TaxID=4081;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Root;

RA Lim E.K., Roberts M.R., Bowles D.J.,

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the annexin family.

CC -1- SIMILARITY: Contains 4 annexin repeats.

DR EMBL; AF079232; AAC97494.1; -.

DR HSSP; Q42657; 1DK5.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR GO; GO:0005444; F:calcium-dependent phospholipid binding; IEA.

DR InterPro; IPR001464; Annexin.

DR InterPro; IPR009118; Annexin\_plant.

DR Pfam; PF00191; Annexin\_4.

DR PRINTS; PR00196; ANNEXIN.

DR PRINTS; PR01814; ANNEXINPLANT.

DR ProDom; PD000143; Annexin\_4.

DR SMART; SM00335; ANX; 4.

DR PROSITE; PS00223; ANNEXIN; 1.

DR KMW Annexin; Calcium; Calcium-binding; Calcium/phospholipid-binding; Repeat.

QY \* SEQUENCE 314 AA; 35801 MW; 18A7A5086A60A6F CRC64;

Query Match 68.0%; Score 1095.5; DB 2; Length 314;

Best Local Similarity 66.1%; Pred. No. 1.6e-61;

Matches 209; Conservative 57; Mismatches 47; Indels 3; Gaps 1;

QY 1 MATLKVSQSDVPAASDADQRLTAFFEGMGNTEDLIISILAHRSABQRKVIROYAHHETGSD 60

DB 1 MASLTVPAAEVPVSAEDCEQLRSAPKGGWINEKLIISILAHRNAQRKILROYAETFGSD 60

QY 61 LKLTLDKELSNDEPERAILLWTLPEGERDALLANEAATKWTSSNOVMEVACTRTSTOLH 120

DB 61 LKELDRRLTHDFEKLVLWTLDPSESDAVLAKEATKWTKSNFVLVEIACRTSRKELVL 120

QY 121 ARQAYHARYKKSLEEDVAHHTTGDPRKLLVSLVTSYRYEGDEVNMTLAKOEAQVHEKIK 180

DB 121 AREAYHARNRKKSLEEDVAHHTTGDPRKLLVPLVSSYRYGDEVDLRLAKAESKVLHEKIS 180

QY 181 DKRYNDEDEVIRILSTRSKAQINATFNRYODDGEELIKSLSEGGDDDKFLALIRSTIOCL 240

DB 181 DKRYSDDEVIRILATRSKAQINATLNHYKDEVEDILKOLE---DDEDFVALLRATIKGL 237

QY 241 TRPELYFVDVLRSAIKKTGDEGALTRIVTTTAEIDLVKVGEEYORNSIPILEKATTKDT 300

DB 238 VYBEHFVEVLRDAINRGTEDEHLLTRVATRAEVDVKIADRYOKRDSIPIGRALAKDT 297

QY 301 RGDYERKMLVALLGEDD 316

Db 298 GGDYENMLVALTGGE 313

Search completed: August 19, 2005, 20:46:32  
Job time : 182 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 07:12:51 ; Search time 4960 Seconds

(without alignments)  
3096.835 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611  
Sequence: 1 MATLKXVSDSPARSDAEQL.....KOTRGDYERKLVALLGEDDA 317

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xjh  
-Q=/cgml1/USPRO.spool/US10690564/runat.19082005.172307.25292/app\_query.fasta\_1.455  
-DB=GenEmbl -QFMT=faatap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptco -NOR=exc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US10690564.@CGML\_1.1.3731.@runat.19082005.172307.25292 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	1611 100.0	954 6	CQ0806050 Sequence
2	1611 100.0	954 8	AF332435 Arabidops
3	1611 100.0	1139 8	BT003359 Arabidops
4	1611 100.0	1192 8	AY086570 Arabidops

5	1611 100.0	1208 8	AY072347 Arabidops
6	1611 100.0	1230 8	AF083913 Arabidops
7	1585 98.4	1159 8	ATUD8415 Arabidops
8	1549.5 96.2	1206 8	ATANNEX Arabidops
9	1507 93.5	954 8	AY356355 Arabidops
10	1450 90.0	8153 8	FL14D7 Arabidops
11	1202.5 74.6	1141 8	GHU073609 Arabidops
12	1193.5 74.1	9448 8	GHU73746 Arabidops
13	1185.5 73.6	2588 6	BD236006 Arabidops
14	1173.5 72.8	1112 8	AF006197 Arabidops
15	1171.5 72.7	1293 6	BD236034 Arabidops
16	1171.5 72.7	1293 6	AR231102 Arabidops
17	1119.5 69.5	1420 8	BT012975 Arabidops
18	1117.5 69.4	1180 8	CANN Arabidops
19	1109.5 68.9	1080 8	CAN130956 Arabidops
20	1107 68.7	1146 8	GHU73747 Arabidops
21	1106.5 68.7	1157 6	AX927134 Arabidops
22	1106.5 68.7	1197 8	AF113545 Arabidops
23	1096.5 68.1	945 8	STU401032 Arabidops
24	1096.5 68.1	1196 8	NTANNE511 Arabidops
25	1095.5 68.0	1160 8	AF079232 Arabidops
26	1093.5 67.9	1086 8	AY351650 Arabidops
27	1087.5 67.5	1172 8	NTANNE671 Arabidops
28	1080 67.0	1242 8	AF079231 Arabidops
29	1067.5 66.3	1182 8	MSANN Arabidops
30	1067 66.2	954 6	CQ0806274 Arabidops
31	1067 66.2	985 8	AY096577 Arabidops
32	1067 66.2	1137 8	AF083914 Arabidops
33	1067 66.2	1160 8	AY085713 Arabidops
34	1067 66.2	1230 8	AY070400 Arabidops
35	1019.5 63.3	1005 8	AY014799 Arabidops
36	1009.5 62.7	1277 8	ZMANNP23 Arabidops
37	1004.5 62.4	1274 8	AK072559 Arabidops
38	991.5 61.5	1120 8	AK175572 Arabidops
39	991.5 61.5	1120 8	AK175641 Arabidops
40	992.5 61.5	1120 8	AK175892 Arabidops
41	982.5 61.0	1003 8	AY014798 Arabidops
42	981.5 60.9	1223 8	AK101787 Arabidops
43	949.5 58.9	1257 8	ZMANNP35 Arabidops
44	879 54.6	997 8	AF188832 Arabidops
45	877 54.4	81494 8	AB019236 Arabidops

## ALIGNMENTS

RESULT 1	CO806050	954 bp	DNA	linear	PAT 10-MAY-2004
LOCUS	CO806050				
DEFINITION	Sequence 2461 from Patent WO2004035798.				
ACCESSION	CO806050				
VERSION	CO806050.1	GI:47111659			
KEYWORDS					
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1 Inze, D., de Veylder, L., and Vlieghe, K.				
AUTHORS	Identification of novel e2f target genes and use thereof				
TITLE	Patent: WO 2004035798-A 2461 29-ARR-2004;				
JOURNAL	CropDesign N.V., (BE)				
FEATURES	Location/Qualifiers				
source	1..954				
ORIGIN	/organism="Arabidopsis thaliana"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:3702"				

Alignment Scores:  
Pred. No.: 8.28e-129  
Score: 1611.00  
Percent Similarity: 100.00%

Length: 954  
Matches: 317  
Conservative: 0

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-690-564-2 (1-317) x C0806050 (1-954)			
QY	1	MetAlaThrLeuLeuValSerAspSerValProAlaProSerAspAlaGluGlnLeu	20
DB	1	ATGGCGACTCTTAAGGTTTCTGATTCCTTCCTGCTCTTCTGATGATGTCGACATTC	60
QY	21	ArgThrAlaPheGluGlyTTPGlyThrAspGluAspLeuLeileSerileLeuAlaHis	40
DB	61	AGAACCGCTTTTAAAGATGGGGTACCAAGACACTGTGATCAATCAACTTGGCTCAC	120
QY	41	ArgSerAlaGluGlnArgLeuValileArgAlaAlaTyrHisGluThrTyrGlyGluAsp	60
DB	121	AGAAAGCTCAACAGAGAAAGTCATCAGCAAGCATCAACAAACTTCGGCGAAGAC	180
QY	61	LeuLeuLeuThrLeuAspLeuGluLeuSerAspPheGluAspAlaileLeuLeuTyr	80
DB	181	CTTCTCAAGACTCTTGAACAGAGCTCTTAACGATTTGAGAGACTATCTTGTGTGG	240
QY	81	ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaSerGluAlaThrLeuAspTyr	100
DB	241	ACTCTGAACCCGAGTGCATGATGCTTTATTTGGCTAATGAACCTCAAAAGATGGACT	300
QY	101	SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis	120
DB	301	TCAAGCAACCAAGTTCTTATGAGAGTTGCTTGAACAGACATCAACGACGCTCTTAC	360
QY	121	AlaArgGlnAlaTyrHisAlaArgTyrLeuSerLeuGluGluAspValAlaHisHis	140
DB	361	GCTAGGCAAGCTTACCATGCTCGCTCAAGAAAGCTCTTGAAGAGAGCGTTGCTCAC	420
QY	141	ThrThrGlyAspPheArgGlyLeuLeuValSerLeuValThrSerTyrArgTyrGluGly	160
DB	421	ACTACCGGACTTCAGAAAGCTTTGGTTTCTTGTTCCTTCAACATCAAGCTCAAGAA	480
QY	161	AspGluValAlaSerMetThrLeuAlaLeuGlnGluAlaLeuLeuValHisGluLeuLeu	180
DB	481	GATGAAGTGAACATGACATTTGGCTTAAGCAAGAAAGCTTACGATGAGAAATCAG	540
QY	181	AspLeuHisTyrAspAspGluAspValileArgileLeuSerThrArgSerLeuAlaGln	200
DB	541	GACAAAGCACTAACATGATGAGATGTTATTAGATCTTGTCCCAAGAAAGCAAGCTCAG	600
QY	201	IleAspAlaThrPheAsnArgTyrGlnAspAspHisGlyGlnGluLeuLeuLeuSerLeu	220
DB	601	ATCAATGCTACTTTTAAACCGTTTCAAGATGATCATGGCAGGAAATTCACAAAGACTT	660
QY	221	GluGluGlyAspAspAspAspPheLeuAlaLeuLeuArgSerThrileGlnCysLeu	240
DB	661	GATGAAGGATATGATGACAAAGTTCTTGCATTTTGAAGTCAACCATTCAGTGTTC	720
QY	241	ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaileAsnTyrThrGlyThr	260
DB	721	ACAAGACCAAGCTTACTTGTTCGATGTTCTTCGTTCAACATTCACAAAGCTGAGACT	780
QY	261	AspGluGluValLeuThrArgileValThrThrArgAlaGluileAspLeuValile	280
DB	781	GATGAAGGACACTCTCAATTAATTTGACCAAGAGCTGAGATTGACTTGAAGGCTCAT	840
QY	281	GlyGluGluTyrGlnArgArgAsnSerileProLeuGluGluValAlaileThrLeuAspTyr	300
DB	841	GGAGAGAGTACCAAGCGCAAGAACAGCATTCCTTTGAGAAAGCTATTACCAAGACACT	900
QY	301	ArgGlyAspTyrGluLeuMetLeuValAlaileLeuGluGlyGluAspAspAla	317
DB	901	CGTGAAGATTACGAGAAAGATGCTCGTGCACCTTCGTGTAAGATGATGTGT	951
RESULT 2			
AF332435	AF332435	954 bp	mRNA linear PLN 07-MAY-2003

DEFINITION	Arabidopsis thaliana clone C00123 (e) putative Ca2+-dependent membrane-binding protein annexin (Atlg35720) mRNA, complete cds.
ACCESSION	AF332435
VERSION	AF332435.1 GI:12083277
KEYWORDS	FLI CDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	1 (bases 1 to 954) Yamada, K., Dale, J.M., Heuan, V.W., Onodera, C.S., Quach, H.L., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A., Arabidopsis Open Reading Frame (ORF) Clones
AUTHORS	Unpublished
TITLE	2 (bases 1 to 954) Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Ertgu, P., Lee, J.M., Toriumi, M., Yu, G., Brooks, S., Chao, Q., Chen, H., Karlin-Neumann, G., Kim, C., Lam, B., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.
REFERENCE	Direct Submission
AUTHORS	Submitted (28-DEC-2000) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
TITLE	3 (bases 1 to 954) Yamada, K., Dale, J.M., Heuan, V.W., Onodera, C.S., Quach, H.L., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.
JOURNAL	Submitted (07-MAY-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT	Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.
FEATURES	Location/Qualifiers
source	1..954 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="1" /clone="C00123 (e)" /ecotype="Columbia" /note="This clone is now in pUNI 51 vector. Previously, it was in pUNI-T3-D/V5-His-10PO under the clone number C00123."
gene	1..954 /gene="Atlg35720"
CDS	1..954 /gene="Atlg35720" /codon_start=1 /evidence="experimental /product="putative Ca2+-dependent membrane-binding protein annexin" /protein_id="AAG48798.1" /db_xref="GI:12083278" /translation="MATIKVSDSVAPSDDAEOLRTAFEGCNGTNDLITSLAHSABQRVTRQAVHYETGDLTKTDKELSNDFERAILMTLEPERGRDLMLNEATKRTSSNOVLMVACTRTSTQLHARQAYHARVYKRSLEEDVAHHTDQFRLLVSLVTSYREGDEVNNTLAKOEAKLVHEKIKDGHVDEIVIRLSTRSAQINATNRKQVODDGEILKSLSEDDDDKFLALRSTQCLTRPELVFVNLRSKINKGTGDEBALTRIVTTREIDLKVIEBYQRNRSIPLEKAITKDTKDGVEKMLVALLGDDA"
ORIGIN	
Alignment Scores:	
Pred. No.:	8,28e-129
Score:	1611.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	8
	Matches: 954
	Conservative: 317
	Mismatches: 0
	Indels: 0
	Gaps: 0

```

US-10-690-564-2 (1-317) x AF332435 (1-954)

QY 1 MetcAthrleuylsValSeraspSerValProAlaProSeraspAapAlaGluGlnLeu 20
Db 1 ATGGCGACTCTTAAGGTTCTGATTCCTGCTCTCTGATGATGCTGAGCAATTG 60
QY 21 ArgThrAlaPheGluGlyTPGlyThAsnGluaspLeuIleSerAlaLeuAlaHis 40
Db 61 AGAACCCCTTTTGAAGGATGGGAGACAGAGACTGATCATATCAATCTTGGCTCAC 120
QY 41 ArgSerAlaGluGlnArgValAlaArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60
Db 121 AGAAGTCTGACAGAGAAAGTATCATGAGCAACATACAGAACTTACGGGAGAC 180
QY 61 LeuLeuLysThrLeuAspLysGluLeuSerAspPheGluArgAlaIleLeuLeuTyr 80
Db 181 CTCTCTCAAGACTCTTGGACAAAGAGCTCTTACCATTTCCAGAGAGCTATCTTGGTGG 240
QY 81 ThrLeuGluProGlyGlyAlaArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrThr 100
Db 241 ACTCTTAAACCCGGTGGAGCTGATGCTTTATGCTATGAGCTACAAAAGATGAGACT 300
QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuHis 120
Db 301 TCAAGCAACCAAGTTCTTATGGAAGTTCCTTGCAACAGACATCAACGCACTGCTTAC 360
QY 121 AlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluAspValAlaHisHis 140
Db 361 GTAGAGCAAGCTTACATGCTCGCTGCTACAGAAAGTCTTGAAGAGAGAGTGGCTCAC 420
QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160
Db 421 ACTACCGGTACTTCAAAAGCTTTGGTTCCTTGTACCTCATACAGTTCAGGAAGA 480
QY 161 AspGluValAlaAsnThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db 481 GATGAAGTGAACATGATGCTGCTACAGCAAGCAAGCTAAGCTGCTCATGAGAAATGAG 540
QY 181 AspLysHisTyrAsnAspGluAspValIleArgGluLeuSerThrArgSerLysAlaGln 200
Db 541 GACAAGACATCAATGATGAGATGTTATTAAGATCTTGTCCACAAAGAACAAAGCTCAG 600
QY 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGluGluGluIleLeuLysSerLeu 220
Db 601 ATCAATGCTACTTTAAACCGTTACCAAGATGATCATGCGAAGAAATTCCAAGAGCTTT 660
QY 221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db 661 GAGGAAGGATATGATGATGACAACTTCTTGCACTTTGAGGTCAACCATTCAGTGTCTG 720
QY 241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
Db 721 ACAAGACCAAGAGCTTCTTGTGATGATGCTTCTGCTTCAACAATCAACAAACTGGAAC 780
QY 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280
Db 781 GATGAAGGAGACCTCATCAATGATGACCAAGACTGATGATGATGATGATGATGATGAT 840
QY 281 GlyGluGlyTyrGlnArgAsnSerIleProLeuGluLysValAlaIleThrLysAspThr 300
Db 841 GAGAGAGAGTACCAAGCGCAAGAACAGCATTCCTTTGAGAAAGCTATTCACCAAGACACT 900
QY 301 ArgGlyAspTyrGlyLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317
Db 901 CGTGAAGATTACGAGAGAGATGCTGCTGCACTTCTCGGTGAAGATGATGCT 951

RESULT 3
LOCUS BT003359 1139 bp mRNA linear PLN 29-JAN-2003
DEFINITION Arabidopsis thaliana Ca2+-dependent membrane-binding protein
anexin (At1g35720) mRNA, complete cds.
ACCESSION BT003359
VERSION BT003359.1 GI:28059005

```

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## PLI CDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 1139)

Southwick, A., Nguyen, M., Tripp, M., Palm, C. J., Jones, T., Wu, T.,

Carninci, P., Chen, H., Cheuk, R., Chan, M. M., Chang, C. H., Dale, J. M.,

Deng, J. M., Hayashizaki, Y., Huan, V. W., Lee, J. M., Ishida, J.,

Kamiya, A., Kawai, J., Kim, C. J., Narusaka, M., Onodera, C. S.,

Quach, H. L., Sakurai, T., Satou, M., Seki, M., Shim, P., Tang, C. C.,

Toroumi, M., Wong, C., Wu, H. C., Yamada, K., Yu, G., Yuan, S.,

Shinozaki, K., Ecker, J., Theologis, A. and Davis, R. W.

Submitted (29-JAN-2003) DNA Sequencing and Technology Center,

Stanford University, 855 California Avenue, Palo Alto, CA 94304,

USA

e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs : Nguyen, M., Tripp, M.,  
Southwick, A., Palm, C. J., Jones, T., Wu, T., Chen, H., Cheuk, R.,  
Chan, M. M., Chang, C. H., Dale, J. M., Deng, J. M., Huan, V. W., Lee, J. M.,  
Kim, C. J., Quach, H. L., Onodera, C. S., Shim, P., Tang, C. C.,  
Toroumi, M., Wong, C., Wu, H. C., Yamada, K., Yu, G., Yuan, S., Ecker, J.,  
Theologis, A. and Davis, R. W.

Nguyen, M., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed  
equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R. W.  
(SSP/Stanford) contributed equally to this work as PIs.

## FEATURES

## source

## Location/Qualifiers

## 1..1139

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="1"

/clone="U17038"

/ecotype="Columbia"

/note="This clone is in DENTR/SD-topo. This is a cloned  
PCR product using RIKEN clone RAFL06-11-P04 (AY072347) as  
a template"

1..1139

/gene="At1g35720"

1..954

/gene="At1g35720"

/codon\_start=1

/product="Ca2+-dependent membrane-binding protein annexin"

/protein\_id="AA029877.1"

/db\_xref="GI:28059006"

/translation="MATLVKSDSVAPSDDAQLRTAFEGWGTNEDLIISILAHRSAR  
ORRVROAHETTYEDLTKLDKELSNDEPRAIILMTLEPGERDLALNEARKRTWS  
NOVLEAVACTRTSTOLHARQAYHARKYKSEDEVAHHTTGPFRKLIVLMTSVRYEG  
DEVNMTAKQEAKEKIKDKGVNDEYIRIISRSKQINATPRRVODDDEETLK  
SIEEGDDDDKFLALRSTIGCTPRELIVDVIRSAINKTGIDEGALTRIVTTTAEID  
LVNIGEEYQRNSIPLEKATVTDGTGDEKMLVALLGEDDA"

## ORIGIN

## Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Beet Local Similarity:	Query Match:	DB:
1	1036-128	1611.00	100.00%	100.00%	8
		Matches:	317		
		Conservative:	0		
		Mismatches:	0		
		Indels:	0		
		Gaps:	0		

US-10-690-564-2 (1-317) x BT003359 (1-1139)

```

Qy      1 MetAlaThrLeuIysValSerAspSerValProlaProSerAspAspAlaGluGlnLeu 20
      1 ATGGCACTCTTAAGGTTTCTGATTCGTCTCCGCTCTCTTCAATGATGCTGACAAATTG 60
Db
Qy      21 ArgThrAlaPheGluGluYTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
      61 AGAACCGCTTTTGAAGGATGGGTACGAACGAGGACTTGATCATCAACTTGGCTCAC 120
Db
Qy      41 ArgSerAlaGluGluIlyrGlyValIleArgGlnAlaSerIleGluThrYrGlyGluAsp 60
      121 AGAAGTGTCTGAACAGAGAAAGTCATCAGGCAAGCAATCAACGAAACCTTACGGCGAAGAC 180
Db
Qy      61 LeuLeuIleThrLeuAspLysGluLeuSerAsnAspPheGluIlyrGlnAlaIleLeuLeuTrp 80
      181 CTTCTCAAGACTCTTGAACAGAGAGCTCTCAACGATTTGAGAGAGACTATCTTGTGTGG 240
Db
Qy      81 ThrLeuGluIlyrProGlyIlyrGluAspAlaLeuLeuAlaAsnGluAlaThrIlyrAspYTrpThr 100
      241 ACTCTGAACCCGGTGGCGGTGATGCTTATGCTTATGCTTAATGAAGCTACAAAAGATGGAGCT 300
Db
Qy      101 SerSerAsnGlnAlaIleuMetGluValAlaCysThrArgThrSerThrGluLeuLeuHis 120
      301 TCAAGCAACCAAGCTTCTTATGAGAGTGTCTTGACAAAGACATCAACGACGCTGCTTAC 360
Db
Qy      121 AlaArgGlnAlaIlyrHisAlaArgYrYrLysSerLeuGluGluAspValAlaHisHis 140
      361 GCTAGGGAAGCTTACCATGCTCGCTTCACAAAGACTCTTGAAGAGACGTTGTCTACACC 420
Db
Qy      141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerYrArgYrGlyGly 160
      421 ACTACCGGTGACTTCAGAAAGCTTTGGTTCTCTTGTACTCATACAGGTACGAAGGA 480
Db
Qy      161 AspGluValaAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
      481 GATGAAGTGAACATGACATTGGCTAAGCAAGAGCTTACGCTGATCAATGAGAAACAAAG 540
Db
Qy      181 AspLysHisIleYrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
      541 GACAGAGACTACATGATGATGAGATGTTATTAGATCTTGTCCACAAAGAGCAAGACTCAG 600
Db
Qy      201 IleAsnAlaThrPheAsnArgYrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
      601 ATCAATGCTACTTTTAAACCGTTTACCAAGATGATCATGCGACAGGAAATTTCTCAAGACTCTT 660
Db
Qy      221 GluGluGluYrAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
      661 GAGGAGAGAGATGATGATGACAGATTCCTTGCACCTTTGAGGTCAACCATTCAGTCTTGG 720
Db
Qy      241 ThrArgProGluLeuYrPheValaAspValaLeuArgSerAlaIleAsnLysThrGlyThr 260
      721 ACAAGACCAAGAGCTTATCTTGTGCATGTTCTTGGTTCAGCATCAACAAACTGGAAGCT 780
Db
Qy      261 AspGluGluYrAlaLeuThrArgIleValaThrThrArgAlaGluIleAspLeuLysValIle 280
      781 GATGAAGAGAGCACTCTCAATATGTCACCAAGAGCTGAGATGACTTGAAGAGTCAAT 840
Db
Qy      281 GluGluGluIlyrGluIlyrArgAsnSerIleProLeuGluLysAlaIleThrIlyrAspThr 300
      841 GGAAGAGAGATGACAGGCGACGAAACAGCATTCCTTTGAGAAACCTTATTAACAAAGACACT 900
Db
Qy      301 ArgGlyAspYrGluLysMetLeuValaLeuLeuLeuGluYrGluLysAspAla 317
      901 CGTGGAGATTACGAGAAAGATGCTGTCGCACTTCTCGGTGAAGATGATGTGT 951
Db

```

```

RESULT 4
LOCUS      AY086570              1192 bp    mRNA           linear   PLN 14-APR-2003
DEFINITION Arabidopsis thaliana clone 25846 mRNA, complete sequence.
ACCESSION  AY086570
VERSION    AY086570.1  GI:21405280
KEYWORDS   Arabidopsis thaliana (thale cress)
SOURCE

```

```

ORGANISM  Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1192)
AUTHORS   Haas, B.-J., Volkovskiy, N., Town, C.D., Troukhan, M., Alexandrov, N.,
           Feldmann, K.A., Flavell, R.B., White, O. and Salzberg, S.L.
TITLE      Full-length messenger RNA sequences greatly improve genome
           annotation
JOURNAL    Genome Biol. 3 (6), RESEARCH0029 (2002)
MEDLINE    22088475
PUBMED     12093376
REFERENCE 2 (bases 1 to 1192)
AUTHORS   Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
           Feldmann, K.
TITLE      Full-length cDNA from Arabidopsis thaliana
JOURNAL    Unpublished
REFERENCE 3 (bases 1 to 1192)
AUTHORS   Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R. and
           Feldmann, K.
TITLE      Direct Submision
JOURNAL    Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
           Malibu, CA 90265, USA
COMMENT    This clone sequence is one of 5,000 Ceres full-length cDNAs made
           available to TIGR and Genbank. The following quality assessment of
           this set was done by comparison with known proteins: two percent of
           the clones are estimated to be 5'-truncated; less than one percent
           are 3'-truncated; approximately two percent represent alternative
           splice variants, including unspliced introns and spliced exons; one
           percent may contain premature stop codons; five percent may have
           frame shifts in a coding region. A sequence is considered to be
           5'-truncated if it lacks the translation initiation start (ATG). A
           sequence is considered to be 3'-truncated if it lacks the
           C-terminal end of the encoded protein. Please note that these cDNA
           sequences are derived from the Ws or Laer ecotypes and therefore
           may contain polymorphisms when compared to sequences from Col-0.
           Geneset carried out the library production and sequencing of the
           full-length clones. Ceres, Inc. carried out the clustering of the
           5' sequences, selection of clones, and sequence assembly.
FEATURES
source     location/Qualifiers
           1..1192
           /organism="Arabidopsis thaliana"
           /mol_type="mRNA"
           /db_xref="taxon:3702"
           /clone="25846"
           36..989
           /codon_start=1
           /product="Ca2+-dependent membrane-binding protein annexin"
           /protein_id="AA063633.1"
           /db_xref="GI:2154612"
           /translation="MATLKVSQVAPSDDAEQLRTAFEGWGTNEDLIISILAHRSAS
           ORKVIROYHYEGRDLKTLDKELSNPEBAIILMTLEPGERDLALNBAATKMTSS
           NOVMEAVACTRSTYGLARQAYARARYKSEEDVAHHTTGDPRLLVLSYVYRG
           DEVNATLAKQEKLVHEKIKDHYNDEIVILILSRSAQINATFNRIQDDHGEITLK
           LSEGGDDDKFLALRSTIQCTREPLKFEVVLKRAINKGTGDEALTRIVTTRAEID
           LKVIIEYQRNSIPLERAKITDTRGDYKMLVALIGBDA"
CDS
Alignment Scores:
Pred. No.:      1..09e-128      Length:      1192
Score:          1611..00        Matches:      317
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    100.00%        Indels:        0
DB:             8               Gaps:         0
US-10-690-564-2 (1-317) x AY086570 (1-1192)
Qy      1 MetAlaThrLeuIysValSerAspSerValProlaProSerAspAspAlaGluGlnLeu 20
      1 ATGGCACTCTTAAGGTTTCTGATTCGTCTCCGCTCTCTTCAATGATGCTGACAAATTG 60
Db
Qy      36 ATGGGCACTCTTAAGGTTTCTGATTCGTCTCCGCTCTTCAATGATGCTGACAAATTG 95
      36 ATGGGCACTCTTAAGGTTTCTGATTCGTCTCCGCTCTTCAATGATGCTGACAAATTG 95
Qy      21 ArgThrAlaPheGluGluYTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40

```



Db		96	AGAACCGCTTTTGAAGGATGGGGGTCAACAGAGCATTTGATCATTCATTGGCTCAC	155
Oy		41	ArgSerAlaGluGlnArgLysValIleArgGlnAlaYrHisGluThrTrpGluAsp	60
Db		156	AGAAGTGCTGAACAGAGGAAGTAGTATCAGGCACAGCATACATGAACAACCTACGGTGAAGAC	215
Oy		61	LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluAspAlaIleLeuLeuTrp	80
Db		216	CTTCCAAAGACTCTTGACAAAGAGGCTCTTAACGACTTCGAGAGGCTAATCTTGTGTGG	275
Oy		81	ThrLeuGluProGlyGluValArgAspAlaLeuLeuAlaAsnGluAlaThrTrpLysArgTrpThr	100
Db		276	ACTCTTGAACCCGGTGAAGCTGATCTTATTATGGCTAATGAAAGCTACAAAAGATTGACT	335
Oy		101	SerSerAsnGlnValLeuMetGluValAlaCythrArgTrpSerThrGlnLeuLeuHis	120
Db		336	TCAAGCAACCAAGTGCTTATGGAATGACTTGCACAGAGCATCAACCAGCTGCTTAC	395
Oy		121	AlaArgGlnAlaTrpYrHisAlaArgTrpLysSerLeuGluLysAspValAlaHisHis	140
Db		396	GCGAGGCAAGCTTACATGCTCTGCATCAAGAAGCTCTTGAAGAAGAGAGTGCTCCACAC	455
Oy		141	ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTrpArgTrpGluGly	160
Db		456	ACTACCGGTGACTTCAGAAAGCTTTGGTCTCTCTTGTTACTCTACACGATGATCAAGAGA	515
Oy		161	AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys	180
Db		516	GATGAAGTGAACATGACATTTGGCTTAAGCAAGAGACTTAAGCTGCTCATGAGAAATCAAG	575
Oy		181	AspLysHisTrpYrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln	200
Db		576	GACAAAGCACTCAATGATGAGAGATGTTATTTGAATCTTGTCACAAAGAAGCAAAGCTCAG	635
Oy		201	IleAsnAlaThrProPheAsnArgTrpGlnAspAspHisGlyGluGlnIleLeuLysSerLeu	220
Db		636	ATCATAGTACTCTTTAACCGTTTACCAAGATGATCAATGGCAGAGAAATCTCAAGAGTCTT	695
Oy		221	GluGluGluYrAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu	240
Db		696	GAGGAGGAGATGATGATGACAAAGTCTCTTGACACTTTTGAGGTCAACATTCAGTGCTTG	755
Oy		241	ThrArgProGluLeuLysTrpYrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr	260
Db		756	ACAAGACCAAGAGCTTACTTGTTCGATGTCTTCGTGAGCATCAACAAACTGSACT	815
Oy		261	AspGluGlyAlaLeuThrArgIleValIleThrThrArgAlaGluIleAspLeuLysValIle	280
Db		816	GATGAAGGAGACTCATCACTGAATGTGTGCCACMAAGCTGAGATTGACTTGAAAGGTCAATT	875
Oy		281	GlyGluGluTrpYrGlnArgArgAsnSerIleProLeuGluLysValaIleThrTrpLysAspThr	300
Db		876	GGAAGGAGTACACAGCGCAGAGAACGCAATTCCTTTGGAGAAAGCTATTATCAAGACACT	935
Oy		301	ArgGlyAspTrpGlyLysMetLeuValaLeuLeuGluGlyGluAspAspAla	317
Db		936	COTGAGATTCAGAGAAGATGCTGCTGCCACTTCTCGGTGAAGATGATGCT	966
RESULT_5				
AY072347				
LOCUS				
DEFINITION			1208 bp mRNA linear PLN 20-JAN-2002	
ACCESSION			Arabisdiopsis thaliana Cas2-dependent membrane-binding protein	
VERSION			anexin1 (Atlg35720) mRNA, complete cds.	
KEYWORDS			AY072347.1 GI:18252242	
SOURCE			FLI CDNA.	
ORGANISM			Arabidopsis thaliana (thale cress)	
			Arabidopsis thaliana	
			Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
			Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;	
			Rosids; euroside II; Brassicales; Brassicaceae; Arabidopsi.	
REFERENCE			1 (bases 1 to 1208)	

**AUTHORS**  
 Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamaya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shim,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.

**TITLE**  
 Direct Submission

**JOURNAL**  
 Submitted (02-JUN-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

**COMMENT**  
 e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RALF cDNAs (RALF cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salik, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RALF cDNAs: Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shim,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

**FEATURES**  
 location/Qualifiers  
 1..1208  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="1"  
 /clone="RALF06-11-P04"  
 /ecotype="Columbia"  
 /note="This clone is in pBluescript"  
 1..1208  
 /gene="Atlg35720"  
 33..986  
 /gene="Atlg35720"  
 /codon\_start=1  
 /product="Ca2+-dependent membrane-binding protein annexin"  
 /protein\_id="P161954.1"  
 /db\_xref="GI:18252243"  
 /translation="MATLKIVSDSPAPDDAOLRTAFEGMGTEIDLIISILHRSAEORVITQAVHRTYGEIDLTKEKLEISNDPFAILLWTLEGEEDALLANATGRWSSNOYLVAVACTRTSTQLHRAQAVYHAKYKSLBEDVAVHTTGDRKLVSIVTSIRLGDEVNMTLAKQBAVLVHKIKDKHYNDSDVRIILSTSKQAINATPKRYQDDHCEILKSLSEGDDDDFFALRLSTIQCLTRPELVPDVARSAINKGTDEGALTRIVTTRAEIDLKVIQEYGRNRNSIPIEKAITKDTDRGDEKMLVALLEGDDA"

**ORIGIN**

Alignment Scores:  
 Pred. No.:  
 Score: 1,11e-128 Length: 1208  
 Percent Similarity: 1611.00 Matches: 317  
 Best local Similarity: 100.00% Conservative: 0  
 Percent local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 8 Gaps: 0

US-10-690-564-2 (1-317) x AY072347 (1-1208)

**QY** 1 MetAlaThrLeuYsValSerAspSerValProAlaProSerAaPaPaLaGlunLeu 20

**Db** 33 ATGGCGAGCTCTTAAGGTTCTGAATCGTTCGTCTCTCTAATGATGCTGAGCAATTG 92

**QY** 21 ArgThrAlaPheGluGlyTTPGjYlThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40

**Db** 93 AGAACCGCTTTTGAAGATGCGGATGACGAAGACGACTGTATATCAATCTTGGCTCAC 152

**QY** 41 ArgSerAlaGluGlnArgValIleValGlnAlaIleTyrHisGluThrTyrGlyGluAsp 60

Db 153 AGAAGTGTGAAAGAGAAAGTCAATCAGGCAAGCATACCAAGAACTTACGGCGAAGAC 212

QY 61 LeuLeuLeuThreThreLeuLeuSerAspAspPheGluArgAlaIleLeuLeuTrp 80

Db 213 CTTCTCAAGACTCTTGAACAAGAGCTCTCTAACGAATTTCAAGAGACTATCTTGGTGTGG 272

QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIlyAspArgTrpThr 100

Db 273 ACTCTTGAACCCCGGTGAGCTGATGCTTTATTTGGCTTAATGAACTTCAAAAAGATGAGACT 332

QY 101 SerSerAsnGluValLeuLeuGluValAlaCysThrArgThrSerThrGluLeuLeuHis 120

Db 333 TCAGGAACCAACAGTTCTTATGGAAGTTGCTTGCAACAAGCAATCAACGCGAGCTGCTTCAAC 392

QY 121 AlaArgGlnAlaIleThiAlaArgTyrIlySerLeuGluGluAspValAlaHisHis 140

Db 393 GCTAGGCAAGCTTACCAATGCTCTGCTCAAGAAAGCTCTTGAAAGAGACGTTGCTCACAC 452

QY 141 ThrThrGlyAspPheArgIlyLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160

Db 453 ACTACCGGTGACTTCAGAAAGCTTTTGCTTCTCTGTTTACCTCATACAGTACGAAGGA 512

QY 161 AspGluValAspMetThrLeuAlaIlyGlnGluAlaIlyLeuValHisGluIlySerIlyLeu 180

Db 513 GATGAAGTGAAACATGACATTGGCTAAGCAAGAAGCTTAAGCTGCTCATGAGAAATCAAG 572

QY 181 AspIlyHisTyrTrpAspArgIlyAspValIleArgIleLeuSerThrTrpSerIlyAlaGln 200

Db 573 GACACAGACTACATAGATGAGATGAGATGTTATTAAGATCTTGCCCAAGAAAGCAAGCTCAG 632

QY 201 IleAspAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuIlySerLeu 220

Db 633 ATCAATGCTACTTTTAAACCTTACCAAGATGATCATGCGCAAGAAATTTCTCAAGACTCTT 692

QY 221 GluGluGlyAspAspAspAspAspIlyPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240

Db 693 GAGGAAGAGATGATGATGACCAAGTCTCTTGCACTTTTGAGTCAACATTCAGTCTTGG 752

QY 241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnIlyThrGlyTrp 260

Db 753 ACAAGACCAAGGCTTATCTTTCGAGTCTTTCGATTCGCAATCAACAAACTGGAAGCT 812

QY 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuIlyValIle 280

Db 813 GATGAAGAGACACTCATCTAGAAATTTGTCACCAAGAGCTGAGTTGACTTGAAGGCTCAT 872

QY 281 GlyGluGluTyrGlnArgArgAsnSerIleProLeuGluIlyAlaIleThrIlyAspTrpThr 300

Db 873 GGAGAGGAGTACCAAGCGCAAGAACAGCATTTCTTTGGAGAAAGCTATTACCAAGACACT 932

QY 301 ArgGlyAspTrpTyrGluIlyMetLeuValAlaLeuLeuGlyGluAspAspAla 317

Db 933 CGTGAAGATTACGAAGAAGATGCTGTCGCACTTCTCGGTGAAGATGATGCT 983

RESULT 6

AF083913 1230 bp mRNA linear PLN 02-JUN-1999

LOCUS AF083913

DEFINITION Arabidopsis thaliana amexin (AmnAcl) mRNA, complete cds.

ACCESSION AF083913.1 GI:4959105

VERSION AF083913.1

KEYWORDS

SOURCE Arabidopsis thaliana (chale cress)

ORGANISM

REFERENCE 1 Arabidopsis thaliana (chale cress); Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.

AUTHORS Clark,G.B. and Roux,S.J.

TITLE Isolation and Characterization of Two Different Arabidopsis Amexin

JOURNAL CDNA (Accession Nos. AF083913 and AF083914) (PGR 99-065)

REFERENCE 2 Plant Physiol. 120 (1), 340 (1999)

AUTHORS Clark,G.B. and Roux,S.J.

TITLE Direct Submission

JOURNAL Submitted (12-AUG-1998) Botany, University of Texas, 24th at

Whitis, Austin, TX 78713, USA

FEATURES

source

1..1230

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/culivar="Columbia"

/db\_xref="taxon:3702"

gene

1..1230

/gene="AmnAcl"

34..987

/gene="AmnAcl"

/note="calcium-binding protein"

/codon\_start=1

/product="amexin"

/protein\_id="AAD34236.1"

/db\_xref="GI:4959106"

/translation="MATLKVSQSDVPAPDDAEQLRTAFEGMGTNEDLIISILAHRSAS

ORKVIROAVHETGSDLKTDKELSNDFERAILMTLEPGERDLANEAATKRTSS

NOVMEVACTRSTQLHARQAVHARVRSKLEEDVAHHTDGFRLIIVTSVYEG

DEVNNTLAKOEAKLVHEKIKDHRNDEIVRILSTRSAQINATRNRYODDGESEILK

SLEEGDDDDKFLALRSTIQCTRELYFVIVLRSAINKGTDEBGALTRIVYTRAEID

LKVIIEYQRRNSIPLEKRAITKDTREYDKMLVALLEGDDA"

ORIGIN

Alignment Scores:

Pred. No.: 1 13e-128 Length: 1230

Score: 1611.00 Matches: 317

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 8 Gaps: 0

US-10-690-564-2 (1-317) x AF083913 (1-1230)

QY 1 MetAlaThrLeuIlyValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20

Db 34 ATGGGCACTCTTAAGGTTTCTGATTCGTGCTGCTCCTTCGATGATGCTGAGCAATTG 93

QY 21 ArgThrAlaPheGluGlyTyrGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40

Db 94 ACAACCGCTTTTGAAGAGATGGGGTACGAACGAGCACTTGATCATATCATCTTGGCTCAC 153

QY 41 ArgSerAlaGluGlnArgIlyValIleArgGlnAlaTyrHisGluThrThrIlyGluAsp 60

Db 154 AGAAGGCTGAACAGAGAAAGTCAATCAGGCAAGCATACCAAGAACTTACGGCGAAGAC 213

QY 61 LeuLeuIlyThrLeuAspIlySerGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80

Db 214 CTTTCTCAAGACTCTTGAACAAGAGCTCTCTTAACGATTTGAGAGAGCTATCTTGTGTGG 273

QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIlyAspArgTrpThr 100

Db 274 ACTCTTGAACCCCGGTGAGCTGATGCTTTATTTGGCTTAATGAACTTCAAAAAGATGACT 333

QY 101 SerSerAsnGluValLeuLeuGluValAlaCysThrArgThrSerThrGluLeuLeuHis 120

Db 334 TCAGGAACCAACAGTTCTTATGGAAGTTGCTTGCAACAAGCAATCAACGAGCTGCTTCAAC 393

QY 121 AlaArgGlnAlaIleThiAlaArgTyrIlySerLeuGluGluAspValAlaHisHis 140

Db 394 GCTAGGCAAGCTTACCATCTGCTGCAAGAAAGTCTCTTGAAGAGAGACTTGTCTCACAC 453

QY 141 ThrThrGlyAspPheArgIlyLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160

Db 454 ACTACCGGTGACTTCAGAAAGCTTTTGCTTCTCTGTTTACCTCATACGATACGAAGGA 513

QY 161 AspGluValAspMetThrLeuAlaIlyGlnGluAlaIlyLeuValHisGluIlySerIlyLeu 180

Db 514 GATGAAGTGAAACATGACATTGGCTAAGCAAGAAGCTTAAGCTGCTCATGAGAAATCAAG 573

QY 181 AspIlyHisTyrTrpAspArgIlyAspValIleArgIleLeuSerThrArgSerIlyAlaGln 200

```

Db      574 GACAGACATCAATGATGAGATGATTTATGATCTTGTCCACAAAGCAAGCTCAG 633
Qy      201 TLeaanaIaThPheanaaTgTgGlnaAspaPhisGlgUgUuileuLySserLeu 220
Db      634 ATCAATGCTACTTTAAACCGTTACCAAGATGATCATGGCAGGAATTTCTCAAGATCTT 693
Qy      221 GUGUGUGUlyAspaAspaAspaAspaAspaAspaAspaAspaAspaAspaAspaAspa 240
Db      694 GAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 753
Qy      241 ThrArgProGluLeuTyrPheValAspValIleuArgSerAlaIleAsnLySThrGlyThr 260
Db      754 ACAAGACACAGAGCTTATCTTTGATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 813
Qy      261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLySValIle 280
Db      814 GATGAGAGACACATCATTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 873
Qy      281 GUGUGUGUlyTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTg 300
Db      874 GAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 933
Qy      301 ArgGlyAspTyrGlyTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTg 317
Db      934 CGTGGAGATTCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984

RESULT 7
ATTU28415      1159 bp      mRNA      linear      PLN 21-OCT-1996
LOCUS          Arabidopsis thaliana annexin-like protein mRNA, complete cds.
DEFINITION
VERSION        U28415.1 GI:1621538
KEYWORDS
SOURCE         Arabidopsis thaliana (chale creess)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosoid II; Brassicales; Brassicaceae; Arabidopsids.
1. (bases 1 to 1159)
Gidrol_X., Sabelki.P.A., Fern.Y.S. and Kush.A.K.
Annexin-like protein from Arabidopsis thaliana rescues delta oxyr
mutant of Escherichia coli from H2O2 stress
Proc. Natl. Acad. Sci. U.S.A. 93 (20), 11268-11273 (1996)
JOURNAL
MEDLINE        9708169
PUBMED         8855345
REFERENCE      2 (bases 1 to 1159)
AUTHORS        Kush.A.K.
TITLE          Direct Submission
JOURNAL        Submitted (06-JUN-1995) Anil K. Kush, Institute of Molecular & Cell
Biology, National University of Singapore, 10 Kent Ridge Crescent,
Singapore 0511, Republic of Singapore
FEATURES
source
1. 1159
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/variety="ecotype Columbia"
/db_xref="taxon:3702"
/clone="poxY5"
/tissue_type="whole seedling"
/clone_lib="lambda UniZapit cDNA library of A.K.Kush"
/dev_stage="4 week-old"
6. .959
/note="annexin-like protein"
/codon_start=1
/protein_id="AC49472.1"
/db_xref="GI:1621538"
/translation="MATTIKSDSVPPSDAEOLRTPFSGKGTNEDELLIISILAHRSAB
QRVIRQAHETYGKDLTKTLKELSNDFERRALLWTLGGEEDALLANEATKRWTS
NOYLMVEACTRTSTOLHARQAYHARYKSLBEDVAHHTGDFRKLIVSLHYSRYEG
DEVNMTLAKOEALVHEKI KDKAHNDEVDVRISTRSKQINATFPRYODDEELIK
SLEGGDDDLPTCTLRSTIOCLRPRIYFQVLRSAINTKGTDEGALVTLVITTRABID
LKVIGREYGRNSIPLEKATTKTRGDYEMLVALLGEDDA"

```

```

ORIGIN
Alignment Scores:
Pred. No.:      1,77e-126      Length:      1159
Score:          1585.00      Matches:     312
Percent Similarity: 98.74%      Conservative: 1
Best Local Similarity: 98.42%      Mismatches: 4
Query Match:    98.39%      Indels:      0
DB:              8      Gaps:          0

US-10-690-564-2 (1-317) x ATTU28415 (1-1159)
Qy      1 MetAlaThrIleuLySValSerAspSerValProAlaProSerAspaAspaIaGluGluLeu 20
Db      6 ATGCGACCTTAAAGATTTCTGATTTCTGTTCTGTTCTGTTCTGTTCTGATATGCTGAGCAATTG 65
Qy      21 ArgThrAlaPheGluGlyTTPGlyThrAngGluAspLeuIleIleSerIleLeuAlaHis 40
Db      66 AGAACCGCTTTGAAGAGATGGGTACGAGACGAGACTTGATCATCATCTTGGCTCAC 125
Qy      41 ArgSerAlaGluGluGlnArgValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60
Db      126 AGAAGTGTGAGACAGAGAAAGTCATCAGGCAAGCATACCAAGAACTTACGGCAAGAAC 185
Qy      61 LeuLeuLySThrIleuAspLySgluLeuSerAsnaAspPheGluArgAlaIleLeuLeuTyr 80
Db      186 CTTCTCAAGACTCTTGAACAAGAGAGCTCTTGAAGATTTGAGAGAGACTATCTTGTTGTGG 245
Qy      81 ThrLeuGluProGlyGluArgAspaAlaLeuLeuAlaAsnGluAlaThrLySArgTTPThr 100
Db      246 ACTCTTGAAACCCGATAGCGTATGCTTTATTTGCTTATGAACTTCAAAAAGATGACT 305
Qy      101 SerSerAsnGlnValIleuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db      306 TCAAGCAACCAAGTTCTTATGAAAGTTGCTTGCAAGAGCATCAACGAGCTGCTTAC 365
Qy      121 AlaArgGlnAlaIleThrIleAlaArgTyrLySlySerLeuGluGluAspValAlaHisHis 140
Db      366 GCTAAGCAAGCTTACCATCTGCTCAACAAGAGTCTTGAAGAGAGCTTGCTCACAC 425
Qy      141 ThrThrGlyAspPheArgLySleuLeuValSerLeuValThrSerTyrArgTyrGlyGly 160
Db      426 ACTACCGGTGACTTCGAAAGCTTTGGTTTCTCTTGTACTTCATACAGTACAGGA 485
Qy      161 AspGluValaAsnMetThrLeuAlaLySglnGluAlaLySLeuValHisGluLySLeu 180
Db      486 GATGAAGTGAACATGATGATGGCTAAGCAAGAGCTTAAGCTGTCATGAGAAATCAAG 545
Qy      181 AspLySHisTyrAsnaAspGluAspValIleArgIleLeuSerThrArgSerLySAlaGln 200
Db      546 GACAACTACTTACATATAGATGATGATGATGATGATGATGATGATGATGATGATGATG 605
Qy      201 TLeaanaIaThPheanaaTgTgGlnaAspaPhisGlgUgUuileuLySserLeu 220
Db      606 ATCAATGCTACTTTTAAACCGTTTACCAAGATGATCATGGCAGGAATTTCTCAAGATCTT 665
Qy      221 GUGUGUGUlyAspaAspaAspaAspaAspaAspaAspaAspaAspaAspaAspaAspa 240
Db      666 GAGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 725
Qy      241 ThrArgProGluLeuTyrPheValAspValIleuArgSerAlaIleAsnLySThrGlyThr 260
Db      726 ACAAGACAGAGCTTATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 785
Qy      261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLySValIle 280
Db      786 GATGAGAGAGCATCTCATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 845
Qy      281 GUGUGUGUlyTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTg 300
Db      846 GAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 905
Qy      301 ArgGlyAspTyrGlyTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTgTg 317

```

Db 906 CGTGAAGATTACGAAGAAGATGCTCGTCGACCTTCGCGTAAGATGATGCT 956  
|||||  
RESULT 8  
LOCUS ATANNEX 1206 bp mRNA linear PLN 13-JUL-1996  
DEFINITION A.thaliana mRNA for annexin.  
ACCESSION X99224  
VERSION X99224.1 GI:1429206  
KEYWORDS annexin; calcium binding protein; cytosolic protein.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1  
AUTHORS Schantz,R., Schantz,M.L. and Houline,G.  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1206)  
AUTHORS Schantz,R.  
TITLE Direct Submission  
JOURNAL Submitted (09-JUL-1996) R. Schantz, Institut de Biologie  
Moléculaire des Plantes, CNRS, 12 Rue de General Zimmer, F- 67084  
Strasbourg, FRANCE  
COMMENT Related sequences: Z47714 & Z33916.  
FEATURES  
source location/Qualifiers  
1..1206  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/sub\_species="ecotype Columbia"  
/db\_xref="taxon:3702"  
/clone="FAFM80"  
/clone\_1b="1ambda Zap II"  
<1..947  
/function="calcium biding protein"  
/codon\_start=3  
/product="annexin"  
/protein\_id="CA67608.1"  
/db\_xref="GI:1429207"  
/db\_xref="GON:Q39001"  
/db\_xref="UniProt/TrEMBL:Q39001"  
/translation="TILKSDSVAPSDDAEQULTAFEGWGTNEDLIISILAHSAEOR  
KVIROAVHETVEGDLTKTLKELSDPEALILMTLEPGRDALLANETKWTSSNQ  
VMEVACTRTSTQLHARQAYHARVYKSLSEEDVAHHTTQDFRLVSLVTSRYEGDE  
VMMTLAKOEKIVHEKIKDKNVDEIVITLSTSAQINATFNRYODDGHSEILKSL  
EEEDDDDDKFLALRSTIOCTTREPILFVIVLRSAIKGTDEBALRIYTTAEIDLK  
VIGSEYQRNSIPLRAITKDNCGDYGMVLVAFGBDA"  
ORIGIN  
Alignment Scores  
Pred. No.: 2,03e-123 length: 1206  
Score: 1549.50 Matches: 307  
Percent Similarity: 98.10% Conservative: 2  
Best Local Similarity: 97.46% Mismatches: 5  
Query Match: 96.18% Indels: 1  
DB: 8 Gaps: 1  
US-10-690-564-2 (1-317) x ATANNEX (1-1206)  
QY 3 ThTleuIyVaIseRAspSerValProAlaProSeRAspAspAlaGluGluLeuAryThr 22  
Db 3 ACTCTTAAGGTTCTGATCTCTGTTCCGCTCTTCGATGATCTGAGCAATTGAGAACC 62  
QY 23 AlaPhGluGluGlyTTPGlyThraENGluAspLeuIleIleSerIleLeuIahIAspSer 42  
Db 63 GCTTTTGAAGAGAGGGGTACGAACGAGACTTATATATCAATCTTGGCTCACAGAAGT 122  
QY 43 AlaGluGlnArgIyValIleArgGlnAlaTyRH;SgluThrTyRgIyGluAspLeuLeu 62  
Db 123 GCTGAACAGAGAAAGTCATCAGGCAAGCTATACCAGAACTTACGGCGAAGCCTTCTC 182  
QY 63 LysThrLeuAspIySgluLeuSerAspPheGluArgAlaIleLeuLeuTTPThrLeu 82  
|||||

Db 183 AAGACTCTTGACAAGAGAGCTCTTAACGATTTGAGAGAGCTATCTTGTGTGACTCTT 242  
QY 83 GluProGlyGluIyAspAlaLeuLeuIahIAspGluIahIleThyIyAspTTPThrSer 102  
Db 243 GAACCCGGGAGACGGATGCTTATTTGGTAAATGAGCTACAAAAGATGAGCTTCAAGC 302  
QY 103 AangInValIleuMetGluValAlaCySerThraRgThrSerThrGlnLeuLeuIahIAsp 122  
Db 303 AACCAAGTCTTATGGAAGTTCCTTGCAACAGACATCAACCACTGCTTCAACCTAGG 362  
QY 123 GlnAlaTyRHAlaIyRgTyRlySeriLeuENGluAspValAlaIahIsthrThr 142  
Db 363 CAAGCTTACCAAGCTCGCTACGAAGAAGTCTTGAAGAGAGAGCTTGCCACCACTACC 422  
QY 143 GlYAspPheArgIyLeuLeuValSerLeuValThrSerTyRgTyRgIyGluIyAspGlu 162  
Db 423 GGTGACTTCAGAAAGCTTTGGTTTCTCTGTTTACTCTATACAGGTACGAAGAATGGA 482  
QY 163 ValAspMetThrLeuAlaIySgIngluAlaIySLeuValIahIleGluIySeriLeuAspIyS 182  
Db 483 GTGAACATGACATTCGCTTACGACAGAGAGTAAAGCTGTCATGAGAAATCAAGACAA 542  
QY 183 HsTyRAspAspGluAspValIleArgIleLeuSerThraRgSerTyRAlaGlnIleAsn 202  
Db 543 CACTCAATGATGAGAGATGTTATTTAGAACTTGTCCACAATMAGCAAGCTCAGATCAAT 602  
QY 203 AlaThrPheAspArgTyRgIyAspAspIahIleGlyGluIleLeuIySerLeuENGlu 222  
Db 603 GCTACTTTTAACCGTTACCAAGATATATCGCGAGAAATTCMAAGTCTTGAAGAA 662  
QY 223 GlYAspAspAspAspIySPhLeuAlaLeuLeuAspSerThrIleGlnCySerThraRg 242  
Db 663 GGAAGATGATGAGACAGAGTCTCTTGCACCTTTGAAGTCAACCATTCAGTGTGACAA 722  
QY 243 ProGluLeuTyRPhyValAspValLeuArgSerAlaIleAsnTyRgIyThraSpGlu 262  
Db 723 CCAGAGCTTACTTGTCCGATGTTCTGCTTCCAGAAATCAACAAAACCTGAATGAA 782  
QY 263 GlYAlaLeuThraRgIleValThrThraRgAlaGluIleAspLeuValIleGlyGlu 282  
Db 783 GAGACCTCACTAGAAATGTGACCAAGAGCTGAGATGACTTGAAGTCAATGAGAG 842  
QY 283 GluTyRgIyAspAspSerIleProLeuGluIyAlaIleThrIySAspThraRgIy 302  
Db 843 GAGTCCAGCGCAGAGACAGACATTCCTTGG---AAGCATTAACCAAGACAAATGTGGA 899  
QY 303 AspTyRgIyLeuMetLeuValAlaLeuLeuGlyGluAspAspAla 317  
Db 900 GATTACGGGAGAGATGCTGTCCCAATTTCCGGGGAAGATGATGCT 944  
|||||  
RESULT 9  
LOCUS AY356355 954 bp mRNA linear PLN 01-SEP-2004  
DEFINITION Brassica juncea annexin mRNA, complete cds.  
ACCESSION AY356355  
VERSION AY356355.1 GI:38049639  
KEYWORDS  
SOURCE Brassica juncea  
ORGANISM Brassica juncea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 954)  
AUTHORS Jami,S.K. and Kirti,P.B.  
TITLE Annexin cDNA from the floral bud of Brassica juncea  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 954)  
AUTHORS Jami,S.K. and Kirti,P.B.  
TITLE Direct Submission  
JOURNAL Submitted (30-JUL-2003) Department of Plant Sciences, School of  
Life Sciences, University of Hyderabad, Gachibowli, Hyderabad,  
Andhra Pradesh 500046, India  
FEATURES  
location/Qualifiers







GRANLEBRMLRQDVALIOPFKLLETPNSLVAMETEKONSVEERFKLFPALDAC  
VQGYVMRKVIVIDGTHLRKRGYGCIVAAASQDANPQVPIAFPIVANSKXDEANTWM

## Alignment Scores:

Pred. No.:	1,236-112	Length:	81513
Score:	1450.00	Matches:	292
Percent Similarity:	90.40%	Conservative:	0
Best Local Similarity:	90.40%	Mismatches:	2
Query Match:	90.01%	Indels:	30
DB:	8	Gaps:	1

US-10-690-564-2 (1-317) x F14D7 (1-81513)

QY	24	Pheglugluytprgylttraspngluaspuleuilelleserileleuallahleargserala	43
DB	6582	TTCCAGAGATGGGTACCAAGAGAGACTGATCATCAATCACTTGGCTCAAGAAAGTCT	6641
QY	44	Gluglunarglyvalillearglnalatyrihsgluthrtyrgllyuaspuleuileuys	63
DB	6642	GAACAGAGAAAGTCATCAGGCAAGCATACCAAGAACTTACGGGAAAGACTTCTCAG	6701
QY	64	ThirleuaplysgluSeraspPheglu-----	74
DB	6702	ACTCTGACAAAGAGCTCTCTAACGATTGCA-GATTGCGTTTCGACCTTTTAAGATC	6760
QY	74	-----	74
DB	6761	GACTTTAAACATCTGTGTGAAGTCATTGTATTCTAATGGTTTTTGGTCATGACAG	6820
QY	75	ArgalaleuleuLeuTPThleuGluproglygluargapalaleuleuAlaasnGlu	94
DB	6821	AGAGCTATCTTGTGTGTGACTTGAACCCGGTACGGTATGCTTTATTTGGCTAATGAA	6880
QY	95	AlathrlybargtTPThrSerSeraspGlnValleuwerGluValAlaCysThrArgThr	114
DB	6881	GCTCAAAAAGATGAGACTTCAAGCAACCAAGTCTTATGAAAGTTGCTTGACAAAGACA	6940
QY	115	SerThrGlnleuLeuHleuAlaArglnAlaArgThiAlaArgTlyGlySerleuGlu	134
DB	6941	TCAAGGAGCTGCTTCACGCTAGGCAAGCTTACATGCTGCTCAAGAAATCTCTTGA	7000
QY	135	GluaspValAlahlesthThrGlyAspPhearglyleuLeuValSerleuValThr	154
DB	7001	GAGACCTTGTCTCACACATCACCGGTGACTTCGAAGCTTTGGTTCTCTTGTACC	7060
QY	155	SerTyArgTyrglygluAspGluValAsnMetThrleuAlalyglnGlnAlalyseu	174
DB	7061	TCATACAGGTACGAAGAGATGAGTGAACATGACATTGGCTAAGCAAGAAAGTAACTG	7120
QY	175	ValHleuGlyvalleyleuaspPheHleuThrAsnaspGluaspValillearglleuSer	194
DB	7121	GTCATATGAAACCAAGGACACACTCATATGATGATGATGATGATGATGATGATGATG	7180
QY	195	ThrArgSerlysalaglnleuaspAlaThrPheAsnArgTyrglnaPaspHleuGlu	214
DB	7181	ACGAAAGCAAGCTTCGATCATATCTTCTTTTAAACGTTTACCAAGATGATGAGGAG	7240
QY	215	GluHleuLeuysSerleuGluGluAlaAspAspAspPheleuAlaleuLeuArg	234
DB	7241	GAATTTCTCAAGAGCTTTGAGGAAGAGATGATGATGATGATGATGATGATGATGATG	7300
QY	235	SerThrilleuGlyleuThrArgPProgluLeuTyThrValaspValleuAspSerAla	254
DB	7301	TCAACCAATTCAGGCTTGAACAGACGAGCTTACTTGTGATGATGATGATGATGATG	7360
QY	255	IlleuAspThrArglyThrAspGluGluValaleuThrArgilleValThrArgAlaGlu	274
DB	7361	ATCAACAAACTGGAATCGATGAGAGCACTCTGGAATTTGACCAACAGAGCTGAG	7420
QY	275	IlleuaspLeuysValilleuGlygluGlyThrArgArgAsnSerilleuLeuGluys	294
DB	7421	ATTGACTTGAAGGTCATTGGAGAGAGATACAGGCGAGAAAGCATTCCTTGGAGAAA	7480

QY 295 AlaleuThrlyaspPThrArglyAspTyrglyuMetleuValAlaleuLeuGlyGlu 314  
DB 7481 GCTATTTCCAAAGAGACTTGTGAGATTTACGAAGATGCTCGTGGCACTTTCGGTGA 7540

QY 315 AspAspAla 317  
DB 7541 GATGATGCT 7549

## RESULT 11

GHU89609

LOCUS GHU89609 1141 bp mRNA linear PLN 30-AUG-1998  
DEFINITION Gossypium hirsutum fiber annexin mRNA, complete cds.  
ACCESSION U89609  
VERSION U89609.1 GI:3493171

## KEYWORDS

SOURCE

ORGANISM

Gossypium hirsutum (upland cotton)  
Gossypium hirsutum  
Bakarae; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

## REFERENCE

1 (bases 1 to 1141)

Shin, H. and Brown, R.M., Jr.

Direct Submission

Submitted (13-FEB-1997)

Botany, The Univ. of Texas at Austin,

Austin, TX 78713, USA

## FEATURES

source

1..1141  
/organism="Gossypium hirsutum"  
/mol\_type="mRNA"  
/strain="Texas Marker1"  
/db\_xref="taxon:3635"  
34..384  
/note="35.5 kda calcium dependent membrane binding protein"  
/codon\_start=1  
/product="fiber annexin"  
/protein\_id="AAC3305.1"  
/db\_xref="GI:3493172"

/translation="MATLVPTTVSVSEDCQRLRAFSQKGTNEGILIDILCHRAE  
ORNLIRKYAETVGEIDLKALDKELSNDEKRLVLAALPAERDALLANEATRWTS  
NOYLMELACTRSANOLHARQVHARVKSLEEDVHHTGDRKLLPLVSSRYEG  
EENMLATBEALHLEKISDKYSDDDYIRLVATRSKQINATLHNYKNEYNDINK  
DKADPKDFLALBSTVCGLVPEYKPVPLAIVNRGTDGALTRVVCRAEYDL  
KIADBYQRNRSVPLTRAIYKTHGDEKLLVLAHVEN"

## ORIGIN

## Alignment Scores:

Pred. No.:	9,66-94	Length:	1141
Score:	1202.50	Matches:	230
Percent Similarity:	85.30%	Conservative:	37
Best Local Similarity:	73.48%	Mismatches:	45
Query Match:	74.64%	Indels:	1
DB:	8	Gaps:	1

US-10-690-564-2 (1-317) x GHU89609 (1-1141)

QY	1	MetAlaThrleuysValSeraspSerValProAlaProSeraspAspAlaGluGlnLeu	20
DB	34	ATGGCCACTTTCACAGTCCCAAGACGATTCCTCGGTGCTGAAGATTTGTAACAGCTA	93
QY	21	ArgThrAlaPhegluGluGlyTTPGlyThrAspGluaspPheleuilelleserileleuAlahle	40
DB	94	AGAAAGCTTTTCAGAGATGGGAACTTAATGAGGCTTAATCATATGATATATTGGGTAC	153
QY	41	ArgSerAlaGluGlnArglyValillearglnAlaTyrihsgluthrtyrgllyuasp	60
DB	154	AGAAAGCCGAGCAAGAAACCTTGAATTCGAAAACTACGCTGAACCTATGAGAGCAT	213
QY	61	LeuLeuysThrleuaspPhegluSeraspPhegluArgAlaleuleuLeuTP	80
DB	214	CTCCTCAAGGCACTTACAAAGAGCTCTCGAATGACTTTGAGAGGCTGTTTGGCTTGG	273
QY	81	ThleuGluProgluGluArgaspAlaleuleuAlaasnGluValaThrlybargtTPThr	100



```

Db      274 GCTCTGATCTCTGCTGAACGTGATGCCCTTTGGCTAATGAAACCAACAAAGGTGAGCT 333
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      101 SerSerAnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuHis 120
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      334 TCAGCAATAGAGGTCTTATGAAATAGCTTGACAGAGGTCTCCAAACCAACGCTTCC 393
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      121 AlArgGlnAlaTyrHisAlaArgTyrLysSerLeuGlnUaAspValAlaHisHis 140
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      394 GCNAGGAGGCTTATCATGCTCTGTATTAAGAAGTCGCTTAAGAAGATGTTGCTCATAC 453
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlnGly 160
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      454 ACCAGTGGCGACCTTCCCTAGCTCCCTCTACCTCTAGTGAGTTCATACATATGAGGGA 513
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysHis 180
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      514 GAGAGGTGAACATGAAATCTGGCGAAAACAGAGCGCAAGTTGCTTCATGAGAAATTTCA 573
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      574 GACCAAGCTTACAGATGATGACGATGTCATAGAGGTTTGGCTACAGAGCAAGGACAG 633
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGlnGluLysLeuLysSerLeu 220
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      634 ATCAATGCAACTCTGATCACTACAAATAATGAAATGCAATGACATTAACCAAGACTTG 693
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      221 GlnGlnGluLysAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      694 ---AAGCTGATCTTAAGATGAGTTCTTGCACTCAAGGTCACAGAGGTCCTTG 750
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      241 ThrArgProGlnLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      751 GTCATCCGGAAGAAAGTATTTTGAGAGGTTCTTCGCTACATCATATGACAGGAGAG 810
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      261 AspGlnGluLysLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle 280
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      811 GATGAAGAGAGCTCTTACTAGAGTTGTTGCATCAGGAGCTGAGGTTGATCTTAAGATCAT 870
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      281 GlnGlnGluLysTyrGlnArgAsnSerIleProLeuGlnLysAlaIleThrLysAspThr 300
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      871 GCAGATGATGACCAAGCAAGAAACAGTCCCACTGATGTCATGTCAAAGGACACT 930
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      301 ArgGlyAspTyrGlnLysMetLeuValAlaLeuLeuGly 313
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      931 CATGAGAGCTATGAAAAATTGCTGCTGCTACTTCGACAGA 969
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
GHU73746      948 bp      mRNA      linear      PLN 24-JAN-2003
LOCUS      Gossypium hirsutum annexin (Annh1) mRNA, partial cds.
DEFINITION      U73746
ACCESSION      U73746.2 GI:15214409
VERSION      U73746.2 GI:15214409
KEYWORDS
SOURCE      Gossypium hirsutum (upland cotton)
ORGANISM      Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eustosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE      1 (bases 1 to 948)
AUTHORS      Delmer,D.P. and Potikha,T.S.
TITLES      Structures and functions of annexins in plants
JOURNAL      Cell; Mol. Life Sci. 53 (6), 546-553 (1997)
MEDLINE      97374485
PubMed      9230934
REFERENCE      2 (bases 1 to 948)
AUTHORS      Potikha,T.S. and Delmer,D.P.
TITLES      Direct Submission
JOURNAL      Submitted (09-OCT-1996) Plant Sciences, Hebrew University, Givat
COMMENT      Ram Campus, Jerusalem 91904, Israel
FEATURES      On Aug 20, 2001 this sequence version replaced gi:1843524.
FEATURES      Location/Qualifiers
source      1..948

```

```

/organism="Gossypium hirsutum"
/mol_type="mRNA"
/strain="Nacala St-2"
/db_xref="taxon:3635"
/tissue_type="cotton fibers"
/dev_stage="21 days post-anthesis"
<1..948
/gene="Annh1"
<1..948
CDS
</gene="Annh1"
/function="evidence suggests that annexins may bind to and
influence activity of callose synthase"
/notes="contains 4 repeats characteristic of other
annexins; shows evidence of conserved calcium-binding
domains; by comparison with other plant annexins as well
as with the size of native cotton fiber annexin, this
clone probably lacks 5' coding sequence for only the
N-terminal Met; calcium binding protein"
/codon_start=1
/product="annexin"
/protein_id="AAB67993.2"
/db_xref="GI:15214410"
/translation="ATLTVPPTVPSVSDCEQLRFAFGKGTNGLIIDLIGHRNAEQ
KULIKTYAETGEDLTKALDKELSENDEKRLVLMALDPAERDALANAEKRTWSN
QVAMEIACRSANQGLHARQVHARVKKSLIEDVNHHTGDFHKLILVSVRYGE
EVMNLTATKALHEKTSNKAYSDDVIVATRSKQINATLHNYKNGENDINKD
LKADKDFEALIRSTKVICVYREKFKVLRALINRGEGALTRVVCRAVDLK
VTADYORRNSVPLTRALVKDTHGDEKLLVLAQVZN"

ORIGIN

Alignment Scores:
Pred. No.:      4.5e-93      Length:      948
Score:          1193.50      Matches:      229
Percent Similarity: 85.26%      Conservative: 37
Best Local Similarity: 73.40%      Mismatches: 45
Query Match:      74.08%      Indels:      1
DB:              8      Gaps:      1

US-10-690-564-2 (1-317) x GHU73746 (1-948)

```

```

Qy      2 AlatrLeuLysValSerAspSerValProAlaProSerAspAspAlaGlnLeuArg 21
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 GCCACTCTTACAGTGGCCAGCAGACGTTCTTCAAGTCTGAGAACTATGTCACATACGA 60
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      22 ThrAlaPheGlnGlyThrGlyThrAnGluAspLeuIleIleSerIleLeuAlaHisArg 41
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61 AAGCCTTTTCAGAGTGGGGAACCTAATAGAGGCTTAATCATATGATATTTGGGTCAAGA 120
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      42 SerAlaGlnGlnArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAspLeu 61
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      121 AATGGGAGCAAGCAAACTGATTCGAAAAAAGCTTCGCGGAAAGCTATGAGAGGATCTC 180
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      62 LeuLysThrLeuAspLysGlnLeuSerAsnAspPheGluArgAlaIleLeuLeuThr 81
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      181 CTCGAAGGCACTAGACAAGAACTTCGATGACTTTGAGAGGCTGTTCTGCTTGGGCT 240
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      82 LeuGlnProGlnGluLysAspAlaLeuLeuAlaAnGlnAlaThrLysArgThrPheSer 101
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      241 CTTGATCTTCGTAACGTGATGCCCTTTTGGTAAATGAAAGCCCAAAAGGTGGACTTCA 300
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      102 SerAnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuHisAla 121
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      301 AGCATCAAGTCTTATGGAATATACCTGCGCAAGAGTCTGCCAACCACTGCTTCACGGA 360
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      122 ArgGlnAlaTyrHisAlaArgTyrLysLysSerLeuGlnGluAspValAlaHisHisThr 141
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      361 AGCGAGGCTTATCATGCTCGTATTAAGAAAGTCTTGAAGAGAGAGTGTCTCATACAG 420
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      142 ThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlnGly 161
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      421 ACTGGGAGCTTCATAGGCTCTCTCACTGAGATCTATGACATATGAGGAGAGG 480
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      162 GluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysHis 181
        |||::|||  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



LOCUS AF006197 1112 bp mRNA linear PLN 22-AUG-2001  
 DEFINITION Lavatera thuringiaca amexin (Amx1) mRNA, complete cds.  
 ACCESSION AF006197  
 VERSION AF006197.1 GI:2459925  
 KEYWORDS  
 SOURCE Lavatera thuringiaca  
 ORGANISM Lavatera thuringiaca  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eucosids II; Malvales; Malvaceae; Malvoideae; Lavatera.  
 REFERENCE 1 (bases 1 to 1112)  
 AUTHORS Vazquez-Tello, A. and Uozumi, T.  
 TITLE Cloning and characterization of a Lavatera thuringiaca cDNA  
 encoding an amexin whose expression is stimulated by low  
 temperature  
 JOURNAL Plant Physiol. (1997) In press  
 REFERENCE 2 (bases 1 to 1112)  
 AUTHORS Breton, G., Vazquez-Tello, A., Danyluk, T. and Sarhan, F.  
 TITLE Two novel intrinsic amexins accumulate in wheat membranes in  
 response to low temperature  
 JOURNAL Plant Cell Physiol. 41 (2), 177-184 (2000)  
 MEDLINE 20255875  
 PUBMED 10795312  
 REFERENCE 3 (bases 1 to 1112)  
 AUTHORS Vazquez-Tello, A. and Uozumi, T.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-MAY-1997) Biological Sciences, Universite du Quebec,  
 Montreal, C.P. 8888, Succ. Centre-Ville, Montreal, Quebec H3C 3P8,  
 Canada

FEATURES  
 source  
 Location/Qualifiers  
 1. .1112  
 /organism="Lavatera thuringiaca"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:61660"  
 /cissue\_type="leaves without petioles"  
 /note="From England"  
 1. .1112  
 /gene="Amx1"  
 2. .952  
 /feature="Amx1"  
 /function="binds acidic phospholipids in a  
 calcium-dependent manner"  
 /note="calcium-dependent, phospholipid binding protein;  
 contains four characteristic repeats found in both animal  
 and plant amexins"  
 /codon\_start=1  
 /product="amexin"  
 /protein\_id="AA571830.1"  
 /db\_xref="GI:2459926"  
 /translation="MATLTVPSTLPVSSEDCQIRKAFSGWGTNEDLIINITLGHNRAD  
 ERNSIRKAYETHEGDELTKALDELSDPERLVILWTLDPPEBDALLANEAATKRTSS  
 NOVIMEIACRSSSDOLIPARQAYHVRKKSLSEDAVHNTGGDRKILLPLVSSRYEG  
 DEYNMTLATREATLIEKISNKRYSDDDPVRILATRSKQINERLHHYNGREVAIDINK  
 DLKADPKDEFLLRSTVKCLVPEKTFEYFELALINKGIDEGALTRVSTAEVDL  
 KIIADBYQRNSVPLTRAIKVDINQDYEKLLVLAGEVA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.83e-91 Length: 1112  
 Score: 1173.50 Matches: 225  
 Percent Similarity: 84.54% Conservative: 43  
 Best Local Similarity: 70.98% Mismatches: 48  
 Query Match: 72.84% Indels: 1  
 DB: 8 Gaps: 1

US-10-690-564-2 (1-317) x AF006197 (1-1112)

QY 1 MetAlaThrLeuIyValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20  
 Db 2 ATGGCTACTCTTACAGTCCCTCCACACCTTCCTGCTGCTGAAGTTGTGAACAAC 61  
 QY 21 ArgThrAlaPheGluGlyTyrGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40

Db 62 AGAAGACCTCTTCAGATGAGGGAAGTAATGAGACCTTAATCATTAATATGGGTCAAC 121  
 QY 41 ArgSerAlaGluGlnArgIyValIleArgGlnAlaTyrHnLeuGluThrTyrGluAsp 60  
 Db 122 CGAATGCGGACGACGAAAGAACTCGATTCGAAAAGCTTATCTGAACCATGAGAAAGAT 181  
 QY 61 LeuLeuIyThrLeuAspIySgluLeuSerAspAspPheGluArgAlaIleLeuLeuTyr 80  
 Db 182 CTCCTCAAGGACCTGGACAAAGAACTCTCAATATGACTTTGAGAGGCTGTGCTTGG 241  
 QY 81 ThrLeuGluProGlyIyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIyAspTyr 100  
 Db 242 ACTCTTGATCTTCCTGAAGGTATGATCATTGTGGAAAAGAACCCAAAGGGGAGCT 301  
 QY 101 SerSerAspGlnValLeuMetGluValAlaCysTrrArgThrSerThrGlnLeuLeuHis 120  
 Db 302 TCAGCAATCAGCTAATTTATGAAATATAGCTTCAGAAAGTTCTTTCGACCACTGCTGCC 361  
 QY 121 AlaArgGlnAlaTyrHisAlaArgTyrIySgluSerLeuGluGluAspValAlaHis 140  
 Db 362 GGAGGCAAGCTTATCATATGCTTATAGAATCGCTTGAAGAGATGTCCTCCATCAC 421  
 QY 141 ThrThrGlyAspPheArgIyLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160  
 Db 422 ACAACTGGGCACTTCCTGAAGCTTCTCCATCTTGTGAAGTTCAATACAGATACAGAGGA 481  
 QY 161 AspGluValAspMetThrLeuAlaIySgluGlnGluAlaIySgluLeuValHisGluIySglu 180  
 Db 482 GATGAAGTAAACATGACTCTGGCGAAACAGAGGCCAAGTTTACTCAGAGAAATCTCA 541  
 QY 181 AspIyHisIyTyrAspAspGluAspIyAspValIleArgIleLeuSerThrArgSerIyAsn 200  
 Db 542 AACAAAGCTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601  
 QY 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGlnIleLeuIySerLeu 220  
 Db 602 ATCAACGAAAGCTTTAATCATCAAAATGATACGCAACTGATGATGATGATGATGATGAT 661  
 QY 221 GluGluGluIyAspAspAspAspIySgluPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
 Db 662 ---MAGCTGACCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 718  
 QY 241 ThrArgProGluLeuIyPheValAspValLeuArgSerAlaIleAsnIyThrGlyThr 260  
 Db 719 GCTACCTCGAAGAAAGATTCGAGAAAGTTCTTCGCTGCAATCAATCAATCAATCAATCA 778  
 QY 261 AspGluGluAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuIyValIle 280  
 Db 779 GATGAAGAGCTCTTAAGAGGTTGTTTCACACAGGCTGAGAGGTTGATCTTAAGATCA 838  
 QY 281 GlyGluGluTyrGlnArgArgAsnSerIleProLeuGluIyValAlaIleThrIyAspTyr 300  
 Db 839 GAGATGATGATCAAGGAAAGAAACAGTCCCACTGATCTGCTGATTTCTCAAGACACT 898  
 QY 301 ArgIyAspTyrGlnIySerLeuValAlaLeuLeuGluIyGluAspAspAla 317  
 Db 899 AATGAGACTACGAAAAATTCGCTGCTGATCTTCTGAGAGGAGTGAAGGCT 949

RESULT 15  
 BD236034 1293 bp DNA linear PART 17-UTL-2003  
 LOCUS BD236034  
 DEFINITION Materials and method for modification of plant cell wall  
 POLYSACCHARIDES.  
 ACCESSION BD236034  
 VERSION BD236034.1 GI:33045804  
 KEYWORDS UP 2002527056-A/56.  
 SOURCE Pinus radiata (Monterey pine)  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.  
 REFERENCE 1 (bases 1 to 1293)  
 AUTHORS Bloksberg, L.N.  
 TITLE Materials and method for modification of plant cell wall

JOURNAL Patent: JP 2002527056-A 56 27-AUG-2002;  
GENESIS RESEARCH AND DEVELOPMENT CORP LTD, FLETCHER CHALLENGE  
FORESTS LTD  
COMMENT OS Pinus radiata (radiata pine)  
PN JP 2002527056-A/56  
PD 27-AUG-2002  
PR 08-OCT-1999 JP 2000575985  
PR 13-OCT-1998 US 09/170862,11-AUG-1999 US 60/148426 PI  
LEONARD NATHAN BLOKBERG  
PC C12N15/09,A01H5/00,C12N5/10,C12N9/00,C12N9/12,C12N15/00,C12N5/  
CC Materials and method for modification of plant cell wall CC  
FH Key location/Qualifiers  
FT source 1..1293 /organism='Pinus radiata (radiata pine)'.  
FT location/Qualifiers  
source 1..1293 /organism='Pinus radiata'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:3347'

ORIGIN

Alignment Scores:  
Pred. No.: 5.05e-91 length: 1293  
Score: 1171.50 Matches: 226  
Percent Similarity: 83.60% Conservative: 39  
Best Local Similarity: 71.29% Mismatches: 51  
Query Match: 72.72% Indels: 1  
DB: 6 Gaps: 1

US-10-690-564-2 (1-317) x BD236034 (1-1293)

QY 1 MetAlaThrLeuLybValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20  
DB 95 ATGTGCACTCTCACACGTCGCGAGCCACTGCCCCCTGTAGCGATGACGAGAGCTC 154  
QY 21 ArgThrAlaPheGluGlyTTPGlyThraAngLysAspLeuLeuLeuLeuLeuLeuLeu 40  
DB 155 CGGACAGCCCTTCCAGAGTGGGAAACAAATGAGAAGCTGATCATATCCATATTTGGGTCAT 214  
QY 41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60  
DB 215 AGGAATGCCGCGCGAGGAAGCTGATTCGGCAAACTATGCCAGACTTACGGCGAGGAC 274  
QY 61 LeuLeuLyThrLeuAspLysGluLeuSerAspAspPheGluArgAlaIleLeuLeuTyr 80  
DB 275 CTCCTCAAGGCAATTGGACAGAGAACTTACCAATGATTTGAGAGGCTGTGCTCTTGG 334  
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTTPhr 100  
DB 335 TCACCTGATCCGGCTGAACGTGATGCTTACTTGGCGGATGAAAGCGAAGAGATGACT 394  
QY 101 SerSerAsnGlnAlaLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
DB 395 TCAAGCAACCAAGTTCCTCATGAAATAGCTTGCACAGGTCTCCGACAGCTTGCTTATG 454  
QY 121 AlaArgGlnAlaTyrHisAlaArgTyrLysLysSerLeuGluGluAspValAlaHisHis 140  
DB 455 GCAAGACAAGCATATCATGCCATACAGAGATCATATGAGAGAGCGTCCCTCACAC 514  
QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlyGly 160  
DB 515 ACAACTGGAGATTGTTGTAAGTTGCTGGTACCTTTGGAGCTTCACTTCACTTATGTGA 574  
QY 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
DB 575 GATGAGGTGAATATGACTTTGGCAAAAGCAGAGGCTTAAGTACTCCACGAGAGATCTCA 634  
QY 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200  
DB 635 GAGAAAGCTTATGCGCATATGATCTCATAGATTTTGGCTACTAGAGCAAAAGCAG 694

QY 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220  
DB 695 GTCAATGCTACGCTGATCATCAAAAATGAGTTTGGAAATGATATCAACAGGATTTG 754  
QY 221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
DB 755 AAAACt---GATCAAAAAGCCGGCTTCTTACTATATCTGAGAGCTACAGTAAAGTCCCTG 811  
QY 241 ThrArgProGluLeuTyrThrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260  
DB 812 ACTCGCCCTGAGAGAAATATTTGAAAAGTTCCTGCTAGCCCATATTAAGCGAGAGACA 871  
QY 261 AspGluGlyAlaLeuThrArgIleValIleThrThrArgAlaGluIleAspLeuLysValIle 280  
DB 872 GATGAGGGGCTCTGACCAAGATGTTGCTTACCAGGCGGAGGTGACATGAAATTATTA 931  
QY 281 GlyGluGluTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300  
DB 932 AGTGAAGATACAGAGAGAGAAATAGCATCCCTCTGATCGTGCATTGTCAAGACACT 991  
QY 301 ArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317  
DB 992 ACTGAGACTATGAAAAAATGCTTCTGCGCATTGATGGCCAGGCT 1042

Search completed: August 22, 2005, 11:53:31  
Job time : 4984 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 06:27:11 ; Search time 665 Seconds

(without alignments)  
2821.892 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611

Sequence: 1 MATLKVSDSPAPSDAEQL.....KDTRGDYKKLVALLGSDDA 317

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seges, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seg length: 0  
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=framed\_p2n.model -DEV=x1h  
-Q=/cgsl\_1/USPTO.spool/US10690564/runat\_19082005\_172306\_25282/app\_query.fasta\_1.455  
-DB=N\_Geneseq\_16Dec04 -OPMT=faastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNIT8=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.cdl  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=p2n -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USR=US10690564.CCGN\_1\_1470.@runat\_19082005\_172306\_25282 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MAPN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_16Dec04: \*  
1: geneseq1980s: \*  
2: geneseq1990s: \*  
3: geneseq2000s: \*  
4: geneseq2001as: \*  
5: geneseq2001bs: \*  
6: geneseq2002as: \*  
7: geneseq2002bs: \*  
8: geneseq2003as: \*  
9: geneseq2003bs: \*  
10: geneseq2003cs: \*  
11: geneseq2003ds: \*  
12: geneseq2004as: \*  
13: geneseq2004bs: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1611	100.0	954	12	ADJ63871 Plant lip
2	1611	100.0	954	12	ADN74566
3	1611	100.0	1192	3	AAC34418 Arabidops
4	1611	100.0	1212	3	AAC34549 Arabidops
5	1566	97.2	1182	3	AAC34339 Arabidops

6	1202.5	74.6	1184	13	ADR63461	Adr63461 Cotton cd
7	1185.5	73.6	2588	3	AAa67100	AAa67100 Eucalyptu
8	1171.5	72.7	1293	3	AAa67128	AAa67128 Pinus rad
9	1171.5	72.7	1293	4	AAa44756	AAa44756 Annexin-1
10	1171.5	72.7	1293	10	ADB94943	ADB94943 Programme
11	1132.5	70.3	1058	10	ADB94854	ADB94854 Programme
12	1106.5	68.7	1157	10	ADP37996	ADP37996 Synchrotr
13	1067	66.2	954	12	ADN74790	ADN74790 Thale cre
14	1067	66.2	1186	3	AAc49728	AAc49728 Arabidops
15	1067	66.2	1160	3	AAc40038	AAc40038 Arabidops
16	1045	64.9	635	6	ABQ66078	ABQ66078 Arabidops
17	1034.5	64.2	951	3	AAc41736	AAc41736 Arabidops
18	988	61.3	1025	3	AAc49738	AAc49738 Arabidops
19	960	59.6	1027	3	AAc37764	AAc37764 Arabidops
20	866.5	53.8	789	3	AAa67031	AAa67031 Pinus rad
21	866.5	53.8	789	3	AAa67098	AAa67098 Eucalyptu
22	866.5	53.8	789	4	AAa44759	AAa44759 Annexin-1
23	866.5	53.8	789	10	ADB94946	ADB94946 Programme
24	793	49.2	589	13	ACN58684	ACN58684 Cotton gy
25	785	48.7	602	13	ACN57334	ACN57334 Cotton gy
26	778.5	48.3	594	12	ADQ07773	ADQ07773 Polynucle
27	775	48.1	591	13	ACN60265	ACN60265 Cotton gy
28	760	47.2	600	13	ACN60840	ACN60840 Cotton gy
29	757.5	47.0	2132	3	AAa67099	AAa67099 Pinus rad
30	742	46.1	618	13	ADR63464	ADR63464 Cotton cd
31	738	45.8	586	13	ACN60745	ACN60745 Cotton gy
32	660	41.0	627	3	AAa67101	AAa67101 Eucalyptu
33	659	40.9	608	13	ACN53624	ACN53624 Cotton an
34	639	39.7	570	13	ACN58229	ACN58229 Cotton gy
35	624.5	38.8	420	3	AAc37667	AAc37667 Arabidops
36	602.5	37.4	966	6	ABZ14750	ABZ14750 Arabidops
37	602.5	37.4	1159	3	AAc45661	AAc45661 Arabidops
38	602.5	37.4	1160	3	AAc42551	AAc42551 Arabidops
39	573	35.6	704	4	AAa67132	AAa67132 Pinus rad
40	573	35.6	704	4	AAa44760	AAa44760 Annexin-1
41	573	35.6	704	10	ADB94947	ADB94947 Programme
42	563	34.9	704	3	AAa67096	AAa67096 Pinus rad
43	520.5	32.3	484	3	AAa67129	AAa67129 Pinus rad
44	520.5	32.3	484	4	AAa44757	AAa44757 Annexin-1
45	520.5	32.3	484	10	ADB94944	ADB94944 Programme

## ALIGNMENTS

RESULT 1	
ID	ADJ63871 standard; DNA; 954 BP.
XX	
AC	ADJ63871:
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Plant lipid metabolism protein OO-10 gene SEQ ID NO:73.
XX	
KW	de; gene; plant; lipid metabolism protein; LMP; seed storage compound;
KW	transgenic plant.
XX	
OS	Unidentified.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	/tag= 1.954
FT	/product= "OO-10"
XX	
XX	WO2004013304-A2.
XX	
PD	12-FEB-2004.
XX	
XX	04-AUG-2003; 2003WO-US024364.
PF	
PR	02-AUG-2002; 2002US-0400803P.
XX	
PA	(BADI ) BASF PLANT SCI GMBH.

XX Miltendorf V, Haertel HA, Bauer J, Oswald O;  
 PI MPI, 2004-157121/15.  
 DR P-PSDB; ADJ63872.  
 XX  
 PT New lipid metabolism proteins and nucleic acids, useful in producing  
 PT transgenic plants with increased levels of seed storage compound, e.g.  
 PT lipid, a fatty acid, a starch or a seed storage protein.  
 XX  
 PS Claim 1, SEQ ID NO 73; 115pp; English.  
 XX  
 CC The invention relates to novel isolated lipid metabolism proteins (LMP)  
 CC and encoding nucleic acids comprising a polynucleotide sequence encoding in  
 CC a polypeptide that functions as a modulator of seed storage compounds in  
 CC a plant. The LMP nucleic acid is useful in producing transgenic plants  
 CC with increased levels of seed storage compound, e.g. lipid, a fatty acid,  
 CC a starch or a seed storage protein, as markers for specific regions of  
 CC the genome and for evolutionary and protein structural studies. The  
 CC present sequence represents an LMP nucleic acid of the invention.

XX Sequence 954 BP; 288 A; 198 C; 229 G; 239 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	2,06e-153	Length:	954
Score:	1611.00	Matches:	317
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-10-690-564-2 (1-317) X ADJ63871 (1-954)

QY 1 MetAlaThrIeuLysValSerAspSerValProAlaProSerAspAlaGluGlnLeu 20  
 DB 1 ATGGCACTCTTAAGGTTCTGATTCCTTCCTTCCTATGATGTCGCAATTG 60  
 QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleSerIleLeuAlaHis 40  
 DB 61 AGAACCGCTTTTAAGGATGGGTACGACAGAGACTTGAATCAATCAATCTTGGCTCAC 120  
 QY 41 ArgSerAlaGluGlnArgValIleArgGlnAlaThrIleGlnThrTyrglyGluAsp 60  
 DB 121 AGAAGGCTGACAGAGAAAGTCATCAGCAGACATCACAGAAACCTACGGCGAAGAC 180  
 QY 61 LeuLeuLysThrIeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80  
 DB 181 CTTCTCAAGACTCTTGAAGAGAGCTCTTAACGATTTTCAGAGACTATCTTGTGTGG 240  
 QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTrpThr 100  
 DB 241 ACTCTTGAAACCCGGTGAAGCTGATGCTTATTGCTTAATCAACTCAAAAAGATGAACT 300  
 QY 101 SerSerAsnGlnValIeuMetGluValAlaCysThrArgThrSerThrGlnLeuHis 120  
 DB 301 TCAAGCAACCAAGTTCTTATGAAAGTTGTTGCAACAAGCATCAACGAGCTTCCAC 360  
 QY 121 AlaArgGlnAlaLysIleAlaArgTyrlsLysSerLeuGluGluAspValAlaHisHis 140  
 DB 361 GCTAGGCAACCTTACATGCTCTCTCAAGAAAGCTCTTGAAGAAGACGTTGCTCACAC 420  
 QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrglyTyrglyGly 160  
 DB 421 ACTACCGGTACTTCAGAAAGCTTTGGTTTCTCTTGTTTCTCAATCAATCAAGTCAAGAA 480  
 QY 161 AspGluValAlaMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
 DB 481 GATGAAGTAAACATGATGCTGCTTAAGCAAGAGCTTGAAGTCCATGAGAAATCAAG 540  
 QY 181 AspLysHisIleTyrsAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200  
 DB 541 GACAAACACTACATGATGATGAGATGTTATTAGATCTTGTCCCAAGAAAGCAAGCTCAG 600

QY 201 IleAsnAlaThrPheAsnArgTyrglnAspAspHisGlyGluGluIleLeuLysSerLeu 220  
 DB 601 ATCAATGCTACTTTTACCGTTTACCAAGATGATCATGGCGAGAAATTTCTCAAGTCTT 660  
 QY 221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
 DB 661 GAGGAAGGAGATGATGATGACAAAGTTCTTGTGACATTTGAGGTCAACCATTCAGTCTTG 720  
 QY 241 ThrArgProGluLeuLeuTyrlsPheValAspValIleuArgSerAlaIleAspLysThrGlyThr 260  
 DB 721 ACAAGACCAAGCTTACTTGTTCATGATGTTCTTCCTCAAGATCAACAAACTGAACT 780  
 QY 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle 280  
 DB 781 GATGAAGAGCACTCACTAGATTTGTGACCAAGAGCTGAGATGACTTGAAGTCAATT 840  
 QY 281 GlyGluGluTyrglnArgArgAsnSerIleProLeuGluLysValIleThrLysAspThr 300  
 DB 841 GGAGAGGAGTACCAAGCAGCAAGACGATTCCTTGGAGAAAGCTATTACCAAGACACT 900  
 QY 301 ArgGlyAspTyrglyLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317  
 DB 901 CGTGAAGATTACGAGAAAGATGCTCGTCCACTTCTCGGTGAAGATGATGCT 951

RESULT 2  
 ADN74566  
 ID ADN74566 standard; cDNA; 954 BP.  
 XX  
 AC ADN74566;  
 XX  
 DT 15-JUL-2004 (first entry)  
 XX  
 DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants SegID 2461.  
 XX  
 KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;  
 KW growth regulator; animal feed product; thale cress;  
 KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.  
 OS Arabidopsis thaliana.  
 XX  
 PN WO2004035798-A2.  
 XX  
 XX 29-APR-2004.  
 PD  
 XX  
 PF 20-OCT-2003; 2003WO-BP011658.  
 XX  
 PR 18-OCT-2002; 2002EP-00079408.  
 XX  
 PA (CROP-) CROPPDESIGN NV.  
 XX  
 PI Inze D, De Veylder L, Vlieghe K;  
 XX  
 DR MPI, 2004-348466/32.  
 DR P-PSDB; ADN74567.  
 XX  
 PT Altering plant characteristics, useful for producing plants for enzyme or  
 PT pharmaceutical production comprises modifying in a plant, expression of  
 PT one or more nucleic acids and/or modifying level or activity of one or  
 PT more proteins.  
 XX  
 PS Claim 1, SEQ ID NO 2461; 134pp; English.  
 XX  
 CC This invention relates to a novel method for altering one or more plant  
 CC characteristics. Specifically, it refers to identifying genes that are up  
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric  
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to  
 CC alter plant characteristics accordingly. The present invention describes  
 CC generating transgenic plants for the production of growth regulators,  
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where  
 CC the altered plant characteristics are selected from increased yield or  
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture  
 CC or physiology, altered endoreplication, biochemistry, signal  
 CC transduction, storage lipid mobilisation and/or altered photosynthesis,



CC each relative to the corresponding wild type plants. Accordingly, these  
CC sequences can also be useful as positive or negative selectable markers  
CC during transformation of cells or tissues. The identified genes play a  
CC role in a variety of biological processes such as DNA replication, cell  
CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as  
CC transcription factors. This polynucleotide sequence is a chalcone synthase  
CC transcribed 1.3 fold or more in plants overexpressing the E2Fa/Dpa  
CC transcription factor, given in an exemplification of the invention.

XX Sequence 954 BP; 288 A; 198 C; 229 G; 239 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	2,06e-153	Length:	954
Score:	1611.00	Matches:	317
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-10-690-564-2 (1-317) x ADN74566 (1-954)

```
QY 1 MetAlaThrLeuValSerAspSerValProAlaProSerAspAlaGluGluLeu 20
Db 1 ATGGCGACTCTTAAGTTCTGATTCCTGCTCTCTCTGATGAGTGAACATTG 60
QY 21 ArgThrAlaPheGluGlyTyrGlyThrAsnGluAspLeuIleSerIleLeuAlaHis 40
Db 61 AGAACCCCTTTTGAAGATGGGTACAGACAGACATTCATATCAATCTGGCTCAC 120
QY 41 ArgSerAlaGluGluGlnArgValIleArgGlnAlaTyrHisGluTyrGlyGluAsp 60
Db 121 AGAAGTCTGACAGAGAAAGTCAATGAGGACACATACACAGAACCTTACGGGAGAC 180
QY 61 LeuLeuValThrLeuAspValGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTyr 80
Db 181 CTCTCTCAAGACTCTTGAAGAGAGCTCTTACAGATTTCCAGAGAGCTATCTTTGG 240
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrValArgTyrThr 100
Db 241 ACTCTTGAACCCGGTGAAGCGTGAATGCTTATGCTTATGAGAGTCAAAAAGATGACT 300
QY 101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGluLeuHis 120
Db 301 TCAAGCAACCAAGTTCTTAAGAAAGTTGCTTGAAGAGACATCAACGACGCTTCAC 360
QY 121 AlaArgGlnAlaTyrHisAlaArgTyrLeuSerLeuGluGluAspValAlaHisHis 140
Db 361 GCTAAGCAAGCTTACCATGCTCGCTACAGAACTCTTGAAGAGACGTTGCTCACAC 420
QY 141 ThrThrGlyAspPheArgGlyLeuLeuValSerLeuValThrSerTyrArgTyrGlyGly 160
Db 421 ACTACCCGTCATTCAGAAAGCTTTGCTTCTTCTTACCTCATACAGGTACGAGAGA 480
QY 161 AspGluValAlaMetThrLeuAlaValGluGluAlaValSerValHisGluValHis 180
Db 481 GATGAAGTGAACATGACATTGGCTAAGCAAGACCTAAGCTGCTCAAGAAATCAAG 540
QY 181 AspValHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerValAlaGln 200
Db 541 GACAGAGCACTACATGATGAGAGATGATTAATTAATCTTGTCCACAAACCAAGCTCAG 600
QY 201 IleAsnAlaThrPheAsnArgTyrGluAspAspHisGlyGluGluIleLeuValSerLeu 220
Db 601 ATCAATCTCTCTTTTAAACCTTACCAAGATGATGCGAGAAATCTTCAAGAGCTCT 660
QY 221 GluGluGluAspAspAspAspValPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db 661 GAGGAAAGAGATGATGAGACAGATCTTCACTTTTGAAGTCAACCATTCAGTGTG 720
QY 241 ThrArgProGluLeuValThrPheValAspValLeuArgSerAlaIleAsnValThrGlyThr 260
Db 721 ACAAGACACAGAGCTTACTTGTGTGATGTTCTTCTTCAACATCAACAAACTGGAAC 780
```

```
QY 261 AspGluGluValaLeuThrArgIleValaThrThrArgAlaGluIleAspLeuValIle 280
Db 781 GATGAAGAGCACTCTCAATTTGATGACCAAGAGCTGAGATTAATTGAAGGTCAAT 840
QY 281 GluGluGluTyrGlnArgArgAsnSerIleProLeuGluValAlaIleThrValAspThr 300
Db 841 GAGAGAGAGTACAGAGCGAGAAACACATCTCTTTGAGAAAGCTTTACCAAGACACT 900
QY 301 ArgGlyAspTyrGluValSerMetLeuValAlaLeuLeuGlyGluAspAspAla 317
Db 901 CGTGAAGATTACAGAAAGATGCTCGTCGCACTTCTCGTGAAGATGATGCT 951

RESULT 3
AAC34418
ID AAC34418 standard; DNA; 1192 BP.
XX
AC AAC34418;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 6580.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132048P.
PR 05-MAY-1999; 99US-0132407P.
PR 06-MAY-1999; 99US-0132484P.
PR 07-MAY-1999; 99US-0132485P.
PR 08-MAY-1999; 99US-0132486P.
PR 09-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134570P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136592P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
```

PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0140091P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0158029P.  
PR 07-OCT-1999; 99US-0158232P.  
PR 08-OCT-1999; 99US-0158330P.  
PR 12-OCT-1999; 99US-0158363P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159584P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160982P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:  
Pred. No. :  
Score:  
Percent Similarity:

2,74e-153  
1611.00  
100.00%

Length: 1192  
Matches: 317  
Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-690-564-2 (1-317) x AAC34418 (1-1192)

```
QY 1 Mec1aThrluLeuValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20
Db 36 ATGGCGACTCTTAAGGTTTCTGATTCTGTTCTGCTCTTCTGATGATGCGACATTG 95
QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
Db 96 AGAACCCCTTTTGAAGATGGGTACAAACGAGACTTGCATCATCAATCTTGCTGCAC 155
QY 41 ArgSerAlaGluGluAspGlyValIleArgGlnAlaThrHisGluThrTrpGlyGluAsp 60
Db 156 AGAAGTCTGAACGAGAAAGATGATCAGCGAACCATACCATGAACCTCAGGGAAGAC 215
QY 61 LeuLeuThrLeuAspLeuGluLeuSerAspPheGluArgAlaIleLeuLeuTrp 80
Db 216 CTTCTCAAGACTCTTGACAAAGACTCTCTACAGACTTCGAGAGACTTACTTGTGCG 275
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIleArgTrpThr 100
Db 276 ACTCTGAACCCGGTGAAGCGTGAATGCTTTATGCTATGAGACTTACAAAAGATGACT 335
QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db 336 TCAGCAACCAAGAGCTTATGAGAGTGTGCAAGAGCATCAGACGAGCTGCTTAC 395
QY 121 AlaArgGlnAlaThrHisAlaArgIleLeuSerLeuGluGluAspValAlaHisHis 140
Db 396 GCGAGGCAAGCTTACCATGCTCGCTACAAAGCTCTTGAAGAGCGTGGCTCACCAC 455
QY 141 ThrThrGlyAspPheArgIleLeuLeuValSerLeuValThrSerTrpArgTrpGlyGly 160
Db 456 ACTACCGGTACTTCAGAAAGCTTTTGTCTCTCTTGTACCTATCAGGTACAGGAAGA 515
QY 161 AspGluValAsnMetThrLeuAlaIleGluGlnAlaIleValLeuValHisGluValIleLeu 180
Db 516 GATGAAGTGAACATGACATGGCTGACAGCAAGAGCTAAGCTGCTCCATGAGAAATCAG 575
QY 181 AspLeuHisTrpAsnAspGluAspValIleArgIleLeuSerThrArgSerIleValGln 200
Db 576 GACAAGACACTACATGATGATGAGATGATTAGATCTTGTCCACAGAACCAAGCTCAG 635
QY 201 IleAsnAlaThrPheAsnArgTrpGlnAspAspHisGlyGluGluIleLeuLeuSerLeu 220
Db 636 ATCAATGCTACTTTTAAACGTTACCAAGATGATCATGGCGAGAAATTTCCAAGATCTT 695
QY 221 GluGluGlyAspAspAspAspAspPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db 696 GAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 755
QY 241 ThrArgProGluLeuThrPheValAspValLeuArgSerAlaIleAsnLeuThrGlyThr 260
Db 756 ACAAGACCAAGAGCTTACTTGTGATGTTCTTGTGATGATGATGATGATGATGATG 815
QY 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280
Db 816 GATGAAGAGACACTCATTAGATTGTGACCAAGACTGATGATGATGATGATGATGATG 875
QY 281 GlyGluGluTrpGlnArgAsnSerIleProLeuGluValAlaIleThrIleAspThr 300
Db 876 GAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 935
QY 301 ArgGlyAspTrpGlyIleMetLeuValAlaLeuLeuGluGluAspAspAla 317
Db 936 CGTGAAGATTACAGAAAGATGCTGCTGCACTTCTCGGTAAAGATGATGCT 986
```

RESULT 4  
AAC34549  
ID AAC34549 standard; DNA; 1212 BP.

```
XX AAC34549;
AC 17-OCT-2000 (first entry)
DT
XX
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 7037.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
OS
XX
PN EP1033405-A2.
PD
XX
XX 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137282P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139494P.
PR 18-JUN-1999; 99US-0139495P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
```

```
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-014187P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 15-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148311P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148564P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149829P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0149930P.

PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 2-8e-153
Score: 1611.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 3
Gaps: 0

US-10-690-564-2 (1-317) x AAC34549 (1-1212)

QY 1 MetAlaThrLeuIyValIserAspSerValPProAlaProSerAspAspAlaGluGluIleu 20
DB 32 ATGGGAGACTCTTAAGGATTCGTGATTCGTCTTCGATGATGCTGAGCAATTG 91
QY 21 ArgThrAlaPheGluGluITripolyThrAsngIuaPleuIleIleSerIleLeuAlaHis 40
DB 92 AGAACCGCTTTTGAAGAGATGGGGTACGAACGAGACTTGATCATATCAATCTTGCTCAC 151
QY 41 ArgSerAlaGluGlnArgIyValIleArgGlnAlaTyRHsGluThrTyrcIyGluuAsp 60
```

```
Db 152 AGAGTCTGTAACAGAGGAAGCATCATGAGCAACATACCTGAACTACGGGAGAC 211
Qy 61 LeuLeuYThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80
Db 212 CTTCTCAAGACTCTTGACAAAGAGCTCTCTTAACACTTCGAGAGCTATCTTGTGTGG 271
Qy 81 ThrLeuGluPyrGlyIuArgAspAlaLeuLeuAlaGluAlaThrIleValArgTrpThr 100
Db 272 ACTCTTAACCCGGTGAAGCGTGAAGCTTATGTGCTATATGAGTACAAAAAGATGACT 331
Qy 101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db 332 TCAGGCAACCAAGTCTTATGGAAGTCTTGACCAAGACATCAAGCGACTCTTCAC 391
Qy 121 AlaArgGlnAlaYrHisAlaArgYrIleYsSerLeuGluAspValAlaHisHis 140
Db 392 GCTAGGCAAGCTTACCATGCTCGCTACAGAAAGTCTTGAAAGAGAGCGTTGCTCACAC 451
Qy 141 ThrThrGluAspPheArgLysLeuLeuValSerLeuValThrSerYrThrArgTrpGluGly 160
Db 452 ACTACCGGTGACTTCAGAAAGCTTTGGTTCTCTTGTACTCATACAGGTACGAAGGA 511
Qy 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db 512 GATGAAGTGAACATGACATTTGGCTTAAGCAAGAACTTAAGCTGCATGAAAAATCAG 571
Qy 181 AspLysHisYrYrAsnAspGluAspValIleArgIleLeuSerThrArgSerYsAlaGln 200
Db 572 GACAAAGCACTTCATGATGAGATGATGATTAATGAACTTTGTCACAAAGAAAGCTCAG 631
Qy 201 IleAsnAlaThrPheAsnArgYrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
Db 632 ATCATATCTACTTTAAACCGTTACCAAGATGATCATGCGGAATTCCTCAAGAGCTT 691
Qy 221 GluGluGluYrAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db 692 GAGGAAGAGATGATGATGACAAAGTCTCTGTTGAGGTCAACATTCAGATGCTTG 751
Qy 241 ThrArgProGluLeuYrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
Db 752 ACAAGACCAAGGCTTACTTGTGATGATGCTCTGTCGACGAATCAACAAACTGCAACT 811
Qy 261 AspGluGluYrAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle 280
Db 812 GATGAAGAGACACTCAGTGAATGTGCCACCAAGAGCTGAGATTGACTTGAAAGTCACT 871
Qy 281 GlyGluGluYrGlnArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
Db 872 GGAGAGAGATGCCAGCGCAAGACAGCATTCCTTGGAAAGCCATTACCAAGACACA 931
Qy 301 ArgGlyAspYrGluYsMetLeuValAlaLeuLeuGlyGluAspAspAla 317
Db 932 CGTGAAGATTACGAAGATGCTGCTGCACCTTCTCGTGAAGATGATGCT 982
RESULT 5
AAC33439
ID AAC33439 standard; DNA; 1182 BP.
XX
AC AAC33439;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 3039.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway; metabolic pathway;
KM promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS
XX EPI033405-A2.
XX
```

```
PD 06-SEP-2000.
XX
PR 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135333P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 21-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140053P.
PR 23-JUN-1999; 99US-0140054P.
PR 24-JUN-1999; 99US-0140695P.
PR 26-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141877P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
```

PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153788P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.

PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160982P.  
PR 25-OCT-1999; 99US-0161040P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 26-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

## Alignment Scores:

Pred. No.: 9 74e-149 Length: 1182  
Score: 1566.00 Matches: 314  
Percent Similarity: 98.74% Conservative: 0  
Best Local Similarity: 98.74% Mismatches: 1  
Query Match: 97.21% Indels: 3  
DB: 3 Gaps: 1

US-10-690-564-2 (1-317) x AAC33439 (1-1182)

QY 1 MetAlaThrLeuIyValSerAspSerValProAlaProSerAspAlaGluGluIleu 20  
DB 33 ATGGGACCTTTAAAGTTTCTGATCTGTCCTCGCTCGATGAGCATTTG 92  
QY 21 ArgThrAlaPheGluGly-TyrGlyThrAsnGluAspLeuIleIleSerIleLeuAlaI 40  
DB 93 AGAACCAGCT-----TTGATGGGGGTACAACGAGGACTTGATCATCATCTTGCTCA 146  
QY 40 sArgSerAlaGluGlnIleArgIleValIleArgIleAlaTyrHisGluIleTyrGlyGluAs 60  
DB 147 CAGAACTCTGAACGAGGAAAGTAGATCAGGCAAGCATACCTGAAAGCTACGGTGAA 206  
QY 60 pleuLeuIySThrLeuAspIyGluIleuSerAsnAspPheGluArgAlaIleLeuIleuTr 80  
DB 207 CCTTCTCAAGACTCTTTGACAAAGAGCTCTTAACGACTTCGAGAGAGCTATCTTGTTG 266  
QY 80 pThrLeuGluIuproGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrTh 100  
DB 267 GACTCTTGAAACCGGGAGACGTGATGCTTATTTAGGCTAATGAAGCTACAAAGAATGAC 326  
QY 100 rSerSerAsnGlnValIleuMetGluValAlaGlyThrArgThrSerThrGluLeuIleuH 120  
DB 327 TTCAGCAACCAAGTGCCTTATGGAAGTAGCTTGACAAAGAGACATCAACGACCTGCTTA 386

```

QY 120 SAIAARGINALATYRHISAIARGLTYRGLYSERLEUGLNUAPVALIAHSHI 140
DB 387 CCGAGGCAAGCTTACATCTCCCTCAAGAACTCTTGAAAGAGACGTGGCTCCCA 446
QY 140 GTHRHGTYASPPHeargylsleuVal1SerleuVal1ThSerTYRArgTYRglu1 160
DB 447 CACTACCGGTGACTTCAGAAAGCTTTGGTCTCTCTGTTACCTACGATACGATGCAAGG 506
QY 160 YASPGIValAsmetThrleuVal1AlvSGINGlVal1AlvleuVal1H1segluyllely 180
DB 507 AGATGAAGTAAACATGATGCTTAAAGCAAGAAAGCTTACCTGATCAGAAACAA 566
QY 180 SAAPLYSHIEYASAsnAspGluAspVal1leArgylleleuSerThraSerlysa1a1 200
DB 567 GGACACACCTACATATATAGATGATTTATTAATCTTGTCCAGAAAGCAAGCTCA 626
QY 200 nleasna1aThrPheAsnArgTYRglAspAspH1seglu1leu1leu1leu1leu1leu1 220
DB 627 GATCAATGCTACTTTTAAACGCTTACCAAGATGATCAGGAGAAATTCACAAAGCTC 686
QY 220 uGIUGlUGLYASpAspAspAspAspAspAspAspAspAspAspAspAspAspAspAsp 240
DB 687 TGAGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 746
QY 240 uThraSPGluLeuTYRPhelVal1AspVal1leuArgSerAl1leu1leu1leu1leu1 260
DB 747 GACAAAGCCAGAGCTTACTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 806
QY 260 RAAPGluGlyAla1leuThraArgylleVal1ThraArgAla1leu1leu1leu1leu1 280
DB 807 TGATGAAGAGCACTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 866
QY 280 eGIUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGlUGl 300
DB 867 TGAGAGAGAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 926
QY 300 RAAGGlyAspTYRglu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1 317
DB 927 TCGTGAAGATTCAGAGAGATGCTCTGTCGACTTCTGAGGAGATGATGCT 978

```

```

XX
DR WPI; 2004-667718/65.
XX
PT New recombinant nucleic acid molecules and polypeptides from Gossypium
PT hirsutum, useful for producing plants with improved biological
PT characteristics (e.g. improved plant cold or drought tolerance).
XX
PS Claim 1; SEQ ID NO 4242; 14pp; English.
XX
CC The invention relates to a recombinant polynucleotide comprising any of
CC the 58798 Cotton plant cDNA sequences mentioned in the specification.
CC Also a recombinant polypeptide comprising any of the 58798 amino acid
CC sequences mentioned in the specification and producing a plant having an
CC improved property. Producing a plant having an improved property
CC comprises transforming a plant with a recombinant construct comprising a
CC promoter region functional in a plant cell operably joined to a
CC polynucleotide comprising a coding sequence for a polypeptide associated
CC with the property, and growing the transformed plant. The polypeptide is
CC useful for improving plant cold tolerance, manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, improving plant
CC drought tolerance, providing increased resistance to plant disease,
CC producing galactomannan (or lignin or plant growth regulators), increasing
CC plant heat tolerance, improving plant tolerance to herbicides, increasing
CC the rate of homologous recombination in plants, improving plant tolerance
CC to extreme osmotic conditions or to pathogens or pests, improving yield
CC by modification of photosynthesis, modifying seed oil or protein yield
CC and/or content, improving yield by modification of carbohydrate, nitrogen
CC or phosphorus use and/or uptake, or improving yield by providing improved
CC plant growth and development under at least one stress condition. The
CC polynucleotide and polypeptide may also be used in recombinant DNA
CC constructs, in physical arrays of molecules, as plant breeding markers,
CC or in computer-based storage and analysis systems. The present sequence
CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20040181830. However only 6585
CC polynucleotide sequences were available, the remaining 52213
CC polynucleotides and all 58798 protein sequences were not present.
XX
SQ Sequence 1184 BP; 375 A; 233 C; 281 G; 295 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.: 6 07e-112 Length: 1184
Score: 1202.50 Matches: 230
Percent Similarity: 85.30% Conservative: 37
Best Local Similarity: 73.48% Mismatches: 45
Query Match: 74.64% Indels: 1
DB: 13 Gaps: 1

```

US-10-690-564-2 (1-317) x ADR63461 (1-1184)

```

QY 1 MetAla1ThrleuVal1SerAspSerVal1ProAlaProSerAspAspAla1leu1leu 20
DB 70 ATGGCCACTTAAAGTGCACAGACAGTTCCTCGGTCTGAGATGTGAAACAGCTA 129
QY 21 ArgThAlaPheGlu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu 40
DB 130 AGAAAGCCCTTTCAGAGATGGGAACTAATGAGGCTTAATCATATATATGGGTAC 189
QY 41 ArgSerAla1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1 60
DB 190 AGAAATGCCAGAGCAAGAACTTGAATTCGAAACCTACGCTGAAACCTATGAGAGGAT 249
QY 61 Leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu 80
DB 250 CTCCTCAAGCACTAAGCAAGAGGCTCTGAAATGACTTTGAGAGGCTGTTTGGCTTGG 309
QY 81 ThrleuGluPArg1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1 100
DB 310 GCTCTTGATCTGCTGAAAGTATGATGCTTTGGCTTAATGAAAGCAACCAAAAGTGGACT 369
QY 101 SerSerAsnGlu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1leu1 120

```



```
Db 370 TCAAGCAATCAGTCTCTTAATGAAATAGCTTGCAAGGCTGCGCAACCACTGCTTAC 429
QY 121 AIAARGIAlaIaTyrHisAlaArgTyrLysSerLeuGluIuAspValAlaHis 140
Db 430 GCAGGAGGCTTATATAGCTCTTAATGAAAGTCCCTGAAAGAGATGTGCTATCAC 489
QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlu 160
Db 490 AGACACTGGGACTTCGTAAGCTCCCTCAACCTCTAGTATGATTCAATACAGATATGAGGA 549
QY 161 AspGluValAsnMetThrLeuAlaLysGluGluAlaLysLeuValHisGluLysIleLys 180
Db 550 GAGAGAGTGAACCTGATCTGGGAAACAGAGCGCAAGTGGCTCATAGAAAATTTC 609
QY 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db 610 GACAAAGCTTACGATATGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
QY 201 ILeAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluLysLeuValIle 220
Db 670 ATCAATGCAACTCTGATCATCTCAAAAATGAATATGAAATGCAATTAACAAGACTTG 729
QY 221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db 730 --AAGCTGATCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
QY 241 ThrArgProGluIuLeuTyrPheValAspValIleuArgSerAlaIleAsnLysThrGlyThr 260
Db 787 GTCATATCCGGAAGAAATTTTGAAGAGTTCTTCGCTGAGCAATCAATGACAGAGAGCG 846
QY 261 AspGluGlyAlaLeuThrArgIleValThrThrArgIleGluIleAspLeuLysValIle 280
Db 847 GATGAAGAGAGCTCTTACTAGATGTTTGTGACTAGAGGCTGAGGTTGATCTAAAGATCATA 906
QY 281 GlyGluGluTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
Db 907 GCAGAGAGATACCAAGGAAAGAAACAGTGTCCCACTACCTGCTGATGTCATGCAAGGACT 966
QY 301 ArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGly 313
Db 967 CATGAGACTATGAAATTTGCTGCTGATCTTGCAAGA 1005

RESULT 7
AAA67100
ID AAA67100 standard; DNA; 2588 BP.
XX
AC AAA67100;
XX
DT 31-OCT-2000 (first entry)
XX
DE Eucalyptus grandis annexin nucleotide sequence SEQ ID NO:28.
XX
KW Eucalyptus grandis; pinus radiata; Monterey pine; modification;
KW plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
KW transgenic plant; ds.
XX
OS Eucalyptus grandis.
XX
PN MO200022092-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999 99WO-N2000169.
XX
PR 13-OCT-1998 98US-00170862.
XX
PR 11-AUG-1999 99US-0148426P.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (LEFT-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Bloksberg LN;
XX
DR WPI; 2000-339328/29.
```

```
DR P-PSDB; AAB16294.
XX
XX New genes encoding proteins involved in a plant polysaccharide
PT biosynthetic pathway, useful for modulating or altering the
PT polysaccharide content, composition or structure of the plant.
XX
XX Claim 1, Page 48-49; 301p; English.
XX
CC The present invention describes isolated polynucleotides (PN) comprising
CC a sequence selected from one of 835 nucleotide sequences given in
CC AAB67073 to AAB67907, their (reverse) complements, sequences producing an
CC expectation (E) value of 0.01 or less compared to the 835 sequences,
CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
CC 835 sequences or sequences that are degenerately equivalent or allelic to
CC the 835 sequences. The polynucleotides are used to modify the activity of
CC a polypeptide involved in a polysaccharide biosynthetic pathway in the
CC plant. They are especially used to modulate or alter the polysaccharide
CC content, composition or structure of the plant. AAB16268 to AAB16340 are
CC proteins encoded by some of the polynucleotide sequence given in the
CC present invention
XX
SQ Sequence 2588 BP; 715 A; 579 C; 628 G; 666 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8,786-110 Length: 2588
Score: 1185.50 Matches: 228
Percent Similarity: 83.91% Conservative: 38
Best Local Similarity: 71.92% Mismatches: 50
Query Match: 73.59% Indels: 1
DB: 3 Gaps: 1

US-10-690-564-2 (1-317) x AAA67100 (1-2588)
QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
Db 1390 ATGTGACTCTACCGTCCCGCAGCCACTGCCCTGTAGCCGATGACTGCGAGAGACTC 1449
QY 21 ArgThrAlaPheGluGlyTyrPglYThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
Db 1450 CGGACAGCTTCGCGACGATGGGAAACAAATGAGAGCTGATCATATTCATATGGGCTCAT 1509
QY 41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60
Db 1510 AGGAATGCGGCGCAGAGAAAGCTGATCGGCAAACTATGCGGAGACTTACCGCAGAGAC 1569
QY 61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLysThr 80
Db 1570 CTCCTCAAGGCAATTGACAGAGAACTTACCAATGATTTGAGAGGCTGCTGCTTTGG 1629
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrThr 100
Db 1630 TCACCTTGAATCCGCGCGAAGCTGATGCTTGGCGAATGAGACGAAAGATGAGACT 1689
QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db 1690 TCAAGCAACCAAGGTTCTTAATGAAATAGCTGCAAGAGTCTCCGACACAGTTCCTTATG 1749
QY 121 AIAARGIAlaIaTyrHisAlaArgTyrLysSerLeuGluIuAspValAlaHis 140
Db 1750 GCAGAGCAAGCAATATCATGATCCGATACAAAGATGCTGTGAAGAGAGAGCTGCTACAC 1809
QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlu 160
Db 1810 ACAAATGAGATTTTCGTAAGTGTGATCCTCTTGTGAGAGTCTTACCTTATATATGGA 1869
QY 161 AspGluValAsnMetThrLeuAlaLysGluGluAlaLysLeuValHisGluLysIleLys 180
Db 1870 GATGAGGTGAATATATGATTTGGCAAAACAGAGGCTTAAGATCTCCACGAGAAATCTCA 1929
QY 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db 1930 GAGAGGCTTATGCGCATAGAGATCTCATAGGATTTGGCTTATGAGAGCAAGACAG 1989
```

[illegible]

Db 932 AGTAGAGATACGAGAGGAGATAGCATCTCTGATCGATTCGATTCAGAGCACT 991  
Qy 301 ArgGlyAspTyrGluValSerMetLeuValAlaLeuLeuGlyGluAspAspAla 317  
Db 992 ACTGGAGACTATGAAAATAATGCTTCTGGCAATGATGGCCAGCTGAGGCT 1042  
RESULT 9  
ID AAF44756 standard; cDNA, 1293 BP.  
XX AAF44756;  
AC AAF44756;  
DT 27-MAR-2001 (first entry)  
XX  
DE Annexin-like protein coding sequence #1.  
XX  
KW Cell death modulator; programmed cell death; PCD; apoptosis;  
KM forestry plant; ss.  
XX  
XX Pinus radiata.  
OS  
PN MO200075331-A1.  
XX  
PD 14-DEC-2000.  
XX  
PF 02-JUN-2000; 2000MO-NZ000086.  
XX  
PR 04-JUN-1999; 99US-00325932.  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
XX  
PI Flinn B, Lasham A;  
XX  
DR MPI; 2001-061724/07.  
DR P-PSDB; AAB5730.  
XX  
PT Novel defender against cell death polynucleotide useful for modulating  
PT programmed cell death pathway and specific development pathways in  
PT forestry plant.  
XX  
XX Claim 1; Page 54-55; 142pp; English.  
XX  
CC The present invention relates to coding sequences (see AAF44740-P44840  
CC and AAF4484-P4484) and proteins (see AAB5714-B65814) involved in  
CC programmed cell death (PCD; apoptosis). The coding sequences and proteins  
CC of the present invention are useful for modulating a PCD or cell death  
CC pathway and various developmental pathways in a forestry plant, by stably  
CC incorporating one of the present coding sequences into the genome of the  
CC forestry plant, where the coding sequence provides a PCD pathway that is  
CC not present in a native form of the forestry plant  
XX  
SQ Sequence 1293 BP; 359 A; 290 C; 313 G; 331 T; 0 U; 0 Other;  
Alignment Scores!  
Pred. No.: 9,34e-109 Length: 1293  
Score: 1171.50 Matches: 226  
Percent Similarity: 83.60% Conservative: 39  
Best Local Similarity: 71.29% Mismatches: 51  
Query Match: 72.72% Indels: 1  
DB: 4 Gaps: 1  
US-10-690-564-2 (1-317) x AAF44756 (1-1293)  
Qy 1 MetaLThreuleuValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20  
Db 95 ATGTCACACTCTCACCCCTCCGACGACCACTGCCCTGTGATCCATATCAGCGACGCTC 154  
Qy 21 ArgThAlaPheGluGlyTyrGlyThrAsnGluAspLeuLeuLeuSerLeuAlaHis 40  
Db 155 CGACACGCTTCCACGAGATGGGGAACAAATGAAAGACTGATCTATATCCATATTTGGGTCA 214  
Qy 41 ArgSerAlaGluGluAlaGlyValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60

Db 215 AGGAATGCGGCGACAGAGAGAGCTGATTCGGCAAACTATGCCGAGACTTACGGCAGAGAC 274  
Qy 61 LeuLeuValThrLeuAspGlySerGluSerLysAsnAspPheGluValAlaIleLeuLeuTyr 80  
Db 275 CTCTCAAGGCAATTTGGACAGAGAACTTACCAATATTTTCAGAGAGGCTGGTGGCTTTGG 334  
Qy 81 ThrLeuGluProGlyGluValArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrThr 100  
Db 335 TCACCTGATCCGGCTGAAACGTATGCTTACCTTGGGAAATGAAAGCCACGAAATGAGACT 394  
Qy 101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGluLeuLeuHis 120  
Db 395 TCAGCAACACAGAGTTCATGAGAAATAGCTTCAGAGAGTCTCCGACAGAGTTCCTATG 454  
Qy 121 AlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluGluAspValAlaHisHis 140  
Db 455 GCAGACACAGCAATATCATGCGCATGCAAGAGTAAAGTAAAGAGAGACGTCCTACACAC 514  
Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160  
Db 515 ACGACTGAGATTTTCTGTAAGTCTGCTGTAACCTCTTGGAGAGCTCTTACCGTAAATGCA 574  
Qy 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
Db 575 GATGAGGTGAATATGACTTTGGCAAAAGAGAGGCTTAAGATCTCCACAGAGAGATCTCA 634  
Qy 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerIleValGln 200  
Db 635 GAGAAAGGCTTAATGGCCATGAGATGCTCATTAAGATTTTGGCTATGAGAGCAACACAG 694  
Qy 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220  
Db 695 GTCAATGCTACGCTGAATCATCAAAATAGATTTGGAATGATATCAACAGAGATTG 754  
Qy 221 GlnGluGlyAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
Db 755 AAAAAGCT--GATCCAAAAGACGCGTCTCTACTATATGAGAGCTACAGTAAAGGCGCTG 811  
Qy 241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260  
Db 812 ACTGCCCTGAGAAAGTATTTTGAAGGTTCTCTCTAGCCATCAATAAAGAGAGCAACA 871  
Qy 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280  
Db 872 GATGAAGGGGCTCTACACAGACTTGTCTACAGAGGCCGAGGTTGACATGAAGTTTATA 931  
Qy 281 GlnGluGluTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300  
Db 932 ACTGAGAGATACCAAGAGAGGAAATGATCCCTTCGATCGTGGCATTTGTCAAGAGACT 991  
Qy 301 ArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317  
Db 992 ACTGGAGACTATGAAAATAATGCTTCTGGCAATGATGGCCAGCTGAGGCT 1042  
RESULT 10  
ID ADB94943 standard; cDNA, 1293 BP.  
XX ADB94943;  
AC ADB94943;  
DT 04-DEC-2003 (first entry)  
XX  
DE Programmed cell death pathway protein annexin cDNA #1.  
XX  
XX programmed cell death; plant development; plant cell cycle; ATT2; DAD1;  
KW Dnaase; 119; 1ad1; nucellin-like aspartic protease; annexin; prohibitin;  
KW fen-like protein; rac2; retinoblastoma-related protein; SIN3; FTRIP;  
KW TEST; xylogenic Rnaase; pur-alpha; cycline protease; RPS-like protein;  
KW gp 91 NADPH oxidase subunit; NFR-like protein; BAG-1;  
KW defender against cell death; lethal leaf spot; lesion stimulating death;  
KW seven in absentia; transcription initiation factor;  
KW testis enhanced gene transcript; gene; ss.

XX Plnus radiata.  
 OS US2003082724-A1.  
 PN  
 XX  
 XX 01-MAY-2003.  
 PD  
 XX 14-AUG-2002; 2002US-00219220.  
 PF  
 XX 04-JUN-1999; 99US-00325932.  
 PR  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 XX  
 XX  
 XX Film B, Lasham A;  
 PI  
 XX WPI; 2003-786916/74.  
 DR  
 XX  
 XX  
 PT New isolated polynucleotide useful for modulating programmed cell death,  
 PT altering the development cycle of plant cells, and subsequently modifying  
 PT plant development.  
 PS  
 PS Claim 1; Page 28-29; 214pp; English.  
 XX  
 CC The invention describes an isolated polynucleotide (1) comprising a  
 CC nucleotide sequence that is one of 145 fully defined sequences of 221-  
 CC 3415 base pairs (bp), given in the specification, and/or its complements,  
 CC reverse complements, reverse sequences, or sequences having 75, 90 or 95  
 CC % sequence identity to or that hybridise under stringent hybridisation  
 CC conditions to one of the 145 sequences. The methods and compositions of  
 CC the present invention to do with (1) are useful for modulating programmed  
 CC cell death and thereby altering the development cycle of plant cells, and  
 CC altering plant development. This sequence encodes a protein associated  
 CC with the programmed cell death pathway.

SQ Sequence 1293 BP; 359 A; 290 C; 313 G; 331 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	9 34e-109	Length:	1293
Score:	1171.50	Matches:	226
Percent Similarity:	83.60%	Conservative:	39
Best Local Similarity:	71.29%	Mismatches:	51
Query Match:	72.72%	Indels:	1
DB:	10	Gaps:	1

US-10-690-564-2 (1-317) x ADB94943 (1-1293)

QY 1 MetAlaThrLeuValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20  
 DB 95 ATGTGCACTCTCACCGTCCGCGACGCACTGCCCTGTAGCGATGACGCGACGCTC 154  
 QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleSerIleLeuAlaHis 40  
 DB 155 CGACACGCTTCCACGATGCGGAAACAATGAGAACGTGATCATCATATTGGGCTCAT 214  
 QY 41 ArgSerAlaGluGlnArgValIleArgGlnAlaThrIleGluThrArgGluAsp 60  
 DB 215 AGGAATGCGCGCGACGAGAAAGCTGATTCGCGAAACCTATGCCAGACTTACGCGAGAC 274  
 QY 61 LeuLeuValThrLeuAspLeuValLeuSerAspPheGluArgAlaIleLeuLeuTrp 80  
 DB 275 CTCTCAAGGCAATTGGACAGAACTTACCAATGATTTCAGAGGCGTGGTCTTTTG 334  
 QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIleValGlyTrpThr 100  
 DB 335 TCACCTGATCCGCGTGAACGTGATGCTACTTGGCGGAAATGAACGAGAAAGATGAGACT 394  
 QY 101 SerSerAsnGlnValLeuMetGlnValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
 DB 395 TCAGGCAACGAGTTCTTCATGAAATAGCCCTGGACGAGGTCTCCGACGACGTTGCTTATG 454  
 QY 121 AlaArgGlnAlaTrpHisAlaArgTrpLeuSerLeuGluGluAspValAlaHisHis 140  
 DB 455 GCAAGACCAAGCATATGATGCCGATACAAAGATGAGTAAGAGACGTCGCTCAACGC 514

QY 141 ThrThrGlyAspPheArgIleLeuValSerLeuValThrSerThrArgTrpGluGly 160  
 DB 515 ACAACCTGAGAAATTTTGTGAAGTTGCTGTGACTCTTGGAGGCTCTTACCGTAATGAG 574  
 QY 161 AspGluValAsnMetThrLeuAlaValGlnGluAlaValLeuValHisGluValIleVal 180  
 DB 575 GATGAGGTGAATTTGACTTTGGCAAAAGCAGAGCTTAAGATATCTCCAGAGAAATCTCA 634  
 QY 181 AspLeuHisIleThrAsnAspGluAspValIleArgIleLeuSerThrArgSerIleValGln 200  
 DB 635 GAGAAAGCTTATGCGCATGAGATCTCATTAAGATTTTGGCTCTAGAGCAAAAGCAG 694  
 QY 201 IleAsnAlaThrThrAsnArgTrpGlnAspAspHisGlyGluGluIleLeuValSerLeu 220  
 DB 695 GTCAATGCTTACGCTGATACCTACAAAAATGAGTTTGAATATGATATCAACAGGATTTG 754  
 QY 221 GlnGluGlyAspAspAspAspLeuPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
 DB 755 AAAACT---GATCAAAAAGACGCGTTCTTACTATATCTAGAGCTACAGTAAAGTCCCTG 811  
 QY 241 ThrArgProGluLeuValThrPheValAspValLeuValSerAlaIleAsnValThrGlyThr 260  
 DB 812 ACTGCGCCCTGAGAAAGTATTTTGAAGTTTCTTCTGCTAGCATCATTAAGCAGAGAAC 871  
 QY 261 AspGluGlyAlaLeuThrArgIleValIleThrThrArgAlaGluIleAspLeuValIle 280  
 DB 872 GATGAAGGGGCTCTGACGAGATGATGCTTACCAAGGCGCAGAGTTGACATGAAGTTTATA 931  
 QY 281 GlyGluGlyIleValThrArgArgAsnSerIleProLeuGluValAlaIleThrIleAspThr 300  
 DB 932 AGTGAAGATGACCAAGGAGGAAATGACATCCCTTGATGTCGTCCATTGTCAAGGACACT 991  
 QY 301 ArgGlyAspTrpGlyIleValMetLeuValAlaLeuLeuGlyGluAspAspAla 317  
 DB 992 ACTGAGACTATGAAAAAATGCTTGTGCAATGATGGCCAGCGTCAGGCT 1042  
 RESULT 11  
 ADB94854  
 ID ADB94854 standard; cDNA; 1058 BP.  
 XX  
 AC ADB94854;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 XX Programmed cell death pathway protein annexin cDNA #6.  
 XX  
 KW KW programmed cell death; plant development; plant cell cycle; ATL2; DAD1;  
 KW Dnaase; 11s; 16s1; nucellin-like aspartic protease; annexin; prohibitin;  
 KW fen-like protein; rac2; retinoblastoma-related protein; SINA; TRITD;  
 KW TEGT; xylogenic RNase; pur-alpha; cyteine protease; RPP5-like protein;  
 KW SP 91 NADPH oxidase subunit; NPR-like protein; BAG-1;  
 KW defender against cell death; lethal leaf spot; lesion stimulating death;  
 KW seven in absentia; transcription initiation factor;  
 KW tests enhanced gene transcript; gene; sr.  
 XX  
 OS Eucalyptus grandis.  
 OS  
 XX  
 PN US2003082724-A1.  
 XX  
 XX  
 PD 01-MAY-2003.  
 XX  
 XX 14-AUG-2002; 2002US-00219220.  
 XX  
 XX  
 PR 04-JUN-1999; 99US-00325932.  
 XX  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 XX  
 XX Film B, Lasham A;  
 PI  
 XX WPI; 2003-786916/74.  
 DR  
 DR P-PDB; ADB94896.  
 XX

PT New isolated polynucleotide useful for modulating programmed cell death,  
 PT altering the development cycle of plant cells, and subsequently modifying  
 PT plant development.

XX Claim 1; Page 138; 21app; English.

XX The invention describes an isolated polynucleotide (I) comprising a  
 CC nucleotide sequence that is one of 145 fully defined sequences of 221-  
 CC 3415 base pairs (bp), given in the specification, and/or its complements,  
 CC reverse complements, or sequences having 75, 90 or 95  
 CC % sequence identity to or that hybridise under stringent hybridisation  
 CC conditions to one of the 145 sequences. The methods and compositions of  
 CC the present invention to do with (I) are useful for modulating programmed  
 CC cell death and thereby altering the development cycle of plant cells, and  
 CC altering plant development. This sequence encodes a protein associated  
 CC with the programmed cell death pathway.

XX Sequence 1058 BP; 306 A; 227 C; 274 G; 251 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	6,39e-105	Length:	1058
Score:	1132.50	Matches:	218
Percent Similarity:	83.91%	Conservative:	48
Best Local Similarity:	68.77%	Mismatches:	50
Query Match:	70.30%	Indels:	1
DB:	10	Gaps:	1

US-10-690-564-2 (1-317) x ADB94854 (1-1058)

QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20  
 Db 72 ATGGCGACTATCCGCGTGCACCGCTCGGTTCCGTCGCCGCTGAGATGCCGAGCTC 131  
 QY 21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40  
 Db 132 CAAAGAGCTTCCGAGATGGAGAACGAATGAAGATCTGATCATCTCACTACGCTCAG 191  
 QY 41 ArgSerAlaGluGluAspValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60  
 Db 192 AGAAAGCCAGCCGCGGAGATGCCGAAAGTATCCGAAACATATCCGAGACATATGGGAGAT 251  
 QY 61 LeuLeuLysThrLeuAspLysGluLeuSerAspAspPheGluGluAlaIleLeuLeuTTP 80  
 Db 252 CTCTCAAGCGCTTGACAAAGAACTCTAGTGACTTGAGAGATCTGCTGTGCG 311  
 QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGlnAlaThrLysArgTTPThr 100  
 Db 312 ACCCTGATCTCGGAGCGTATGATCTTGTCCAAATGAAGCTCAAGAGATTGACT 371  
 QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
 Db 372 TCAGACACTGGGGTTCTCATGAAATTCCTTGACAGAGCTTCATATGAGATTATTCAG 431  
 QY 121 AlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluGluAspValAlaHisHis 140  
 Db 432 GTGAGCGAGGCTATCATGCTCGTTATGAAGAACTCTTGAAAGAAACATCCGATATCAC 491  
 QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160  
 Db 492 ACTACTGGGATTTCCGCAAGCTGCTTGTCCCTGCAAGATVACCTTCGATATGAGGG 551  
 QY 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
 Db 552 CCGAGAGTGAACATGACATTCGCAAGATCAGAGGCTTAAGTACTTCAATGAAAGATTCC 611  
 QY 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200  
 Db 612 GAGAAAGCTTACATCATGTAGCTCATCAGAAATTGTACTCAAGAAAGTAAAGCTCAG 671  
 QY 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluLysLeuLysSerLeu 220  
 Db 672 CTTAATGCAACCTCTCAATTAATCAACAAATGAGTTGGGAATGCCATCAACAGAGATTCTG 731

QY 221 GluGluGlyAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
 Db 732 ---MAGCGTANCCAAATGATCAATTCTGAAACTGCTGAGATCGCAATTAAGGCTTG 788  
 QY 241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260  
 Db 789 ACTTATCCAGAGAAATCACTTGTGAGAAAGTACGCTCGGCCATCAACAGCTGGAAACA 848  
 QY 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280  
 Db 849 GATGATGGGCTCTTACCAAGATACCACTCGTGGAGAGTGAATGACATGGAGGCAATC 908  
 QY 281 GlyGluGlyTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300  
 Db 909 AAAGAAAGTACCAAGAAAGAAAGAGTGTCTCTAGATCGGCTATTGCAATGACACA 968  
 QY 301 ArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317  
 Db 969 TCTGGGGACTACAAAGAGATGCTCTCGCTTGTGTGACATGAGAGATGCT 1019

RESULT 12  
 ADF37996  
 ID ADF37996 standard; cDNA; 1157 BP.  
 AC ADF37996;  
 XX 12-FEB-2004 (first entry)  
 DT  
 XX Synchronised tobacco BY2 cDNA sequence SEQ ID NO:55.  
 DE  
 XX Identification; validation; plant; agrochemical; gene; ss.  
 KW  
 OS Nicotiana tabacum.  
 XX  
 EN W02003085115-A2.  
 XX  
 PD 16-OCT-2003.  
 XX  
 PF 08-APR-2003; 2003WO-BP003703.  
 XX  
 PR 10-APR-2002; 2002EP-00447062.  
 XX  
 PR 15-JUL-2002; 2002US-0396124P.  
 XX  
 PA (CROP-) CROPDSEIGN NY.  
 XX  
 PI Inze D, Broekaert W;  
 XX  
 DR WPI; 2003-804308/75.  
 XX  
 PT Identifying and validating plant genes or proteins as targets for  
 PT agrochemicals, useful for producing agrochemical-resistant plants,  
 PT comprises determining and down regulating the gene or protein expression  
 PT profiles of a plant.  
 XX  
 XX Claim 12; SEQ ID NO 55; 183bp; English.  
 PS  
 CC The present invention describes a method for identifying and validating  
 CC plant genes/proteins as targets for agrochemicals comprising determining  
 CC the gene or protein expression profiles of a plant and downregulating the  
 CC expression of the gene or protein in the plant or plant cell. Also  
 CC described: (1) methods for screening candidate agrochemical compounds,  
 CC comprising the use of the above method or the use of any of the 785 fully  
 CC defined nucleotide sequences (ADP37942 to ADF38726) or protein sequences,  
 CC or their homologues, functional fragments or derivatives; (2) a method  
 CC for producing an agrochemical resistant plant, comprising the use of the  
 CC above-mentioned nucleotide or protein sequences; (3) an isolated nucleic  
 CC acid that is identified by any of the above methods or that comprises at  
 CC least a part of a nucleic acid sequence chosen from any of the 785  
 CC nucleotide sequences given in the specification; (4) a plant tolerant to  
 CC an agrochemical, in which the expression level of one or more of the  
 CC nucleic acid sequences given in the specification is modulated; and (5) a  
 CC harvestable part of the plant described above. The method is useful in  
 CC identifying and validating plant targets for agrochemicals or in

CC producing an agrochemical resistant plant. The nucleic acid or protein  
 CC can be used as a target for an agrochemical compound, particularly  
 CC herbicide. The present sequence represents a synchronised tobacco BY2  
 CC cDNA nucleotide sequence which is used in the exemplification of the  
 CC present invention.

XX Sequence 1157 BP; 362 A; 213 C; 270 G; 312 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	3.07e-102	Length:	1157
Score:	1106.50	Matches:	214
Percent Similarity:	81.07%	Conservative:	43
Best Local Similarity:	67.51%	Mismatches:	59
Query Match:	68.68%	Indels:	1
DB:	10	Gaps:	1

US-10-690-564-2 (1-317) x ADF37996 (1-1157)

```

Qy 1 MetAlaThrLeuValSerAspSerValProLalProSerAspAspAlaGluGluLeu 20
Db 45 ATGGCGAGCTTAAAGTCCAAACATCTGTTCCAGAACCTTATGAGATGGCTGAGCAATC 104
Qy 21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuLeuSerIleLeuAlaHis 40
Db 105 AAAAAGCTTTTGTCTGATGGGGTACAAATGAGGCACTTATTATTCAGATTCTGGGCAT 164
Qy 41 ArgSerAlaGluGlnAlaGlyValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60
Db 165 AGAATGACACACACCAACCAAGTTATCCGAAACCTTATCTGCAGCTTATGAGAGAT 224
Qy 61 LeuLeuThrLeuAspLeuValSerAspAspPheGluValAlaIleLeuLeuTyr 80
Db 225 CTCTCTCAAGACTTGGATGCTGAACCTGACAAAGTATTTTCAAGCTGAGCTTCTGTGG 284
Qy 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrThr 100
Db 285 ACTTTGAGTCTCTGTGAGCGGACGCGCTACTTGGTTATGAGAGTACCAACGCTGACT 344
Qy 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db 345 TCTAGCAATGGGTATCTTGGAAATGCTTGTACAAAGGCTTCTGTGATGATCTTTAAG 404
Qy 121 AlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluGluAspValAlaHisHis 140
Db 405 GCGAGGAGGCTTACACAGCTCGATACAAAGAACCTTAAAGAAAGTGTGCTTATCAC 464
Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160
Db 465 ACAACTGGGATTTCCGTAAAGCTTTGGTTCCTTTTAATCGCATTCAGATACGAAGCA 524
Qy 161 AspGluValAsnMetThrLeuAlaLysGlnAlaLysLeuValHisGluLysIleLys 180
Db 525 GAAGAGCGCAACATGACATTCGCAAGAAAGAGCAAAATACTACACAGCAAGAAATCTCT 584
Qy 181 AspLysHisThrAspAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db 585 GACAAAGGCTTACATGATGAGAGGCTCATCCGAATTATTTCTACTAGAGTAAAGCAG 644
Qy 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
Db 645 CTGAATGCAACATTCACACCTACCTTGCACCAACATGGCAGTGAATCAACAAAGATCTG 704
Qy 221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db 705 GAAACT---GATTCGTGATGATGATGATGAAATTAATCAAGCCACAGCAATAGATGCTTG 761
Qy 241 ThrArgProGluLeuValTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
Db 762 AAAAACCAGAGAAACCTTTGAGAAAGTCTTGATGGCTTATCAAGGGTACAGGACACA 821
Qy 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280
Db 822 GACGAATGGGACCTTACTATGATGTTGTCACTACCTCGGCTGAATTTGACATGAAACGATTC 881

```

```

Qy 281 GlyGluGlyTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
Db 882 AAAGAAGAGTACCATGAGAGAAACAGTGTCCATTGACCGTCAATTGCTGAGACACT 941
Qy 301 ArgGlyAspTyrGlnLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317
Db 942 TCAGAGACTATGAAAGATGCTTCTGCTTTGATTTGGGATGAGATGCT 992

```

# RESULT 13

ADN74790  
 ID ADN74790 standard; cDNA; 954 BP.

XX  
 AC ADN74790;

XX  
 DT 15-JUL-2004 (first entry)

XX  
 DE Thale cress cDNA repressed in E2Fa/Dpa expressing plants SeqID 2655.

XX  
 KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;

XX  
 KW growth regulator; animal feed product; thale cress;

XX  
 KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.

XX  
 OS Arabidopsis thaliana.

XX  
 PN W02004035798-A2.

XX  
 PD 23-APR-2004.

XX  
 PF 20-OCT-2003; 2003WO-BE011658.

XX  
 PR 18-OCT-2002; 2002EP-00079408.

XX  
 PA (CROP-) CROPDESIGN NV.

XX  
 PI Inze D, De Veylder L, Vlieghe K;

XX  
 DR WPI; 2004-348466/32.

XX  
 DR P-PADB; ADN74791.

XX  
 PT Altering plant characteristics, useful for producing plants for enzyme or

XX  
 PT pharmaceutical production comprises modifying in a plant, expression of

XX  
 PT one or more nucleic acids and/or modifying level or activity of one or

XX  
 PT more proteins.

XX  
 PS Claim 1; SEQ ID NO 2685; 134bp; English.

XX  
 CC This invention relates to a novel method for altering one or more plant

XX  
 CC characteristics. Specifically, it refers to identifying genes that are up

XX  
 CC - or down-regulated in transgenic plants overexpressing the heterodimeric

XX  
 CC E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to

XX  
 CC alter plant characteristics accordingly. The present invention describes

XX  
 CC generating transgenic plants for the production of growth regulators,

XX  
 CC enzymes, therapeutics, pharmaceuticals and animal feed products, where

XX  
 CC the altered plant characteristics are selected from increased yield or

XX  
 CC biomass, enhanced survival capacity, stress tolerance, plant architecture

XX  
 CC or physiology, altered endoreplication, biochemistry, signal

XX  
 CC transduction, storage lipid mobilisation and/or altered photosynthesis,

XX  
 CC each relative to the corresponding wild type plants. Accordingly, these

XX  
 CC sequences can also be useful as positive or negative selectable markers

XX  
 CC during transformation of cells or tissues. The identified genes play a

XX  
 CC role in a variety of biological processes such as DNA replication, cell

XX  
 CC wall biosynthesis, nitrogen and/or carbon metabolism or they function as

XX  
 CC transcription factors. This polynucleotide sequence is thale cress cDNA

XX  
 CC repressed 1.3 fold or more in plants overexpressing the E2Fa/Dpa

XX  
 CC transcription factor, given in an exemplification of the invention.

XX  
 SQ Sequence 954 BP; 301 A; 216 C; 216 G; 221 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	2.38e-98	Length:	954
Score:	1067.00 <td>Matches:</td> <td>203</td>	Matches:	203
Percent Similarity:	79.81% <td>Conservative:</td> <td>50</td>	Conservative:	50



Best Local Similarity: 64.04% Mismatches: 64  
Query Match: 66.23% Indels: 0  
DB: 12 Gaps: 0  
US-10-690-564-2 (1-317) x ADNT4790 (1-954)

```
QY      1 MetAlaThrLeuIysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 ATGGCGCTCTCAAAAGTCCCAAGCAATGTTCTCTCCGAAATATACGCCGAGCAACTC 60
QY      21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61 CACAGAGCTTTTCAGGATGGGTACCAAGAGAGCTGATCATCAATACATAGCTCAC 120
QY      41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaLeuHisGluThrTrpGlyLysAsp 60
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      121 AGAAGCGACACACAGCAGCTTGATCCGAGCGTTTATGACGTAACCTACAAATGAGAT 180
QY      61 LeuLeuIleThrLeuAspLysGluLeuSerAsnAspPheGluAlaIleLeuLeuTrp 80
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      181 CTTCTCAAGCATTTAGACAAAGAGCTTTCTAGGACCTTTGAGAGACTGTGATGTGG 240
QY      81 ThrLeuGluProGlyGlyLysAspAlaLeuLeuAlaAsnGluAlaThrLysArgTrpThr 100
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      241 ACTCTGATCCACAGAGAGATGCTTATTTGGCTTAAAGAAATCCACCAAGATGTTCAACC 300
QY      101 SerSerAsnGlnValIleuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      301 AAGAGCAATTTGGGTCTTGTAATGCTTGACAAAGGCTGCTCTTGAGCTTATCAAG 360
QY      121 AlaArgGlnAlaIleThrIleAlaArgTrpLysSerLeuGluGlnAspValAlaHisHis 140
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      361 GTCAGCAAGCTTACCAAGCTGCATACAAAGAAATCATAGAGAAATGTCCGCAACAC 420
QY      141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValIleThrSerTrpArgTrpGly 160
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      421 ACATCTGGTCACTTGTGAAGCTTGTGCTTCTCTGCTGAGCACTTTCAGTATGAAGA 480
QY      161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      481 GATGATGTGAACATGATGCTTGAGAGTCTGAGACTGACGTAAGTACTTCCAGAGAGCTCA 540
QY      181 AspLysHisIleTrpAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      541 GAGAAATCTTACAGTGCATGATGCTTCATCAGAAATCTTGACAAACAAAGAAAGCAACG 600
QY      201 IleAsnAlaThrPheAsnArgTrpGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      601 CTGGTGCAACACTCAACCACTCAACCAAGAGTATGAAACGCCCATTAACAGAACTTG 660
QY      221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      661 AAGGAAGATCGGACGACATGACTACATGAACCTACAGCTGTAAATCACATGTTG 720
QY      241 ThrArgProGluLeuTrpPheValAspValIleuArgSerAlaIleAsnLysThrGlyThr 260
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      721 ACATACCTCGAAGACATTTTGAAAGGTTCTTCGCTCATCATCAACAAATGGAACA 780
QY      261 AspGluGlyAlaLeuThrArgIleValIleThrArgAlaGluIleAspLeuLysValIle 280
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      781 GACGAATGGGAGACTTAACCCGATCGTACACGAACTGAATGACATGGAACCCATC 840
QY      281 GlyGluGluTrpGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      841 AAAGAGATATATACGAGAAACAGCATTCCTTTGACCGTGTCTATTCGCCAAGACACT 900
QY      301 ArgGlyAspTrpGlyLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      901 TCTGGTGACTATGAGGACATGCTTGTGCTCTCTGGAACATGGCATGCT 951
```

RESULT 14  
AAC49728  
ID AAC49728 standard; DNA; 1156 BP.

```
XX      XX AAC49728;
AC      AC 18-OCT-2000 (first entry)
XX      XX
DT      DT
DE      DE Arabidopsis thaliana DNA fragment SEQ ID NO: 62219.
XX      XX Hybridisation assay; genetic mapping; gene expression control;
KW      KW protein identification; signal transduction pathway; metabolic pathway;
XX      XX promoter; termination sequence; ss.
XX      XX Arabidopsis thaliana.
OS      OS
PN      PN EP1033405-A2.
XX      XX
PD      PD 06-SEP-2000.
XX      XX
XX      XX
PF      PF 25-FEB-2000; 2000EP-00301439.
XX      XX
PR      PR 25-FEB-1999; 99US-0121825P.
PR      PR 05-MAR-1999; 99US-0123180P.
PR      PR 09-MAR-1999; 99US-0122548P.
PR      PR 23-MAR-1999; 99US-0125788P.
PR      PR 25-MAR-1999; 99US-0126264P.
PR      PR 29-MAR-1999; 99US-0126785P.
PR      PR 01-APR-1999; 99US-0127462P.
PR      PR 06-APR-1999; 99US-0128234P.
PR      PR 08-APR-1999; 99US-0128714P.
PR      PR 16-APR-1999; 99US-0129845P.
PR      PR 19-APR-1999; 99US-0130077P.
PR      PR 21-APR-1999; 99US-0130449P.
PR      PR 23-APR-1999; 99US-0130510P.
PR      PR 28-APR-1999; 99US-0130891P.
PR      PR 30-APR-1999; 99US-0131449P.
PR      PR 30-APR-1999; 99US-0132048P.
PR      PR 30-APR-1999; 99US-0132407P.
PR      PR 04-MAY-1999; 99US-0132484P.
PR      PR 05-MAY-1999; 99US-0132485P.
PR      PR 06-MAY-1999; 99US-0132486P.
PR      PR 07-MAY-1999; 99US-0132487P.
PR      PR 11-MAY-1999; 99US-0132863P.
PR      PR 14-MAY-1999; 99US-0134256P.
PR      PR 14-MAY-1999; 99US-0134218P.
PR      PR 14-MAY-1999; 99US-0134219P.
PR      PR 14-MAY-1999; 99US-0134221P.
PR      PR 14-MAY-1999; 99US-0134370P.
PR      PR 18-MAY-1999; 99US-0134768P.
PR      PR 19-MAY-1999; 99US-0134941P.
PR      PR 20-MAY-1999; 99US-0135124P.
PR      PR 21-MAY-1999; 99US-0135353P.
PR      PR 24-MAY-1999; 99US-0135629P.
PR      PR 25-MAY-1999; 99US-0136021P.
PR      PR 27-MAY-1999; 99US-0136392P.
PR      PR 28-MAY-1999; 99US-0136782P.
PR      PR 01-JUN-1999; 99US-0137222P.
PR      PR 03-JUN-1999; 99US-0137528P.
PR      PR 04-JUN-1999; 99US-0137502P.
PR      PR 07-JUN-1999; 99US-0137724P.
PR      PR 08-JUN-1999; 99US-0138094P.
PR      PR 10-JUN-1999; 99US-0138540P.
PR      PR 10-JUN-1999; 99US-0138847P.
PR      PR 14-JUN-1999; 99US-0139119P.
PR      PR 16-JUN-1999; 99US-0139452P.
PR      PR 16-JUN-1999; 99US-0139453P.
PR      PR 17-JUN-1999; 99US-0139492P.
PR      PR 18-JUN-1999; 99US-0139454P.
PR      PR 18-JUN-1999; 99US-0139455P.
PR      PR 18-JUN-1999; 99US-0139456P.
PR      PR 18-JUN-1999; 99US-0139457P.
PR      PR 18-JUN-1999; 99US-0139458P.
PR      PR 18-JUN-1999; 99US-0139459P.
PR      PR 18-JUN-1999; 99US-0139460P.
PR      PR 18-JUN-1999; 99US-0139461P.
```



```
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140685P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144864P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147319P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.

PR 25-AUG-1999; 99US-0150566P.
PR 25-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161509P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161820P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 3.05e-98
Score: 1067.00
Percent Similarity: 79.81%
Best Local Similarity: 64.04%
Query Match: 66.23%
DB: 3
Gaps: 0

US-10-690-564-2 (1-317) x AAC49728 (1-1156)

QY 1 MetAlaThreIleuValSerAseValProAlaProSerAsePheAlaGluGlnLeu 20
   |||||:::|||||
Db 95 ATGGCTCTCTCAAAATCCCAAGCAATGTTCTCTCCGAGATACCCGACACTC 154
   |||||
QY 21 ArgThrAlaPheGluGlyTyrPglYThrAngIuaPleuIleIleSerIleLeuAlaHis 40
   |||||
Db 155 CACAAGCTTTTTCAGAGTGGGGTACCAAGAAAGCTGATCATATCACTAGCTCAC 214
   |||||
QY 41 ArgSerAlaGluGlnArgValIleArgGlnAlaTyrHisGluThrTyrGlyGluAer 60
```

```
Db 215 AGAAGCGACGACAGCAGCTTGATCCGACGGCTTATTCAGCTACCTACATGAGAT 274
Qy 61 LeuLeuLysThrLeuLysThrLeuLysSerAsnAspPheGluArgAlaLeuLeuTrp 80
Db 275 CTTCTCAAGACATTAGACAAAGAGCTTTGAGGACCTTGAGAGACCTGGATGTGTGG 334
Qy 81 ThrLeuGluProGluArgAlaLeuLeuAlaAsnGluAlaThrLysArgTrpTrp 100
Db 335 ACTCTTGATCCACAGAGAGATGCTTATTTGGCTTAAGAACCCAGAGATGTCACC 394
Qy 101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGluLeuHis 120
Db 395 AAGAACATGGGCTTCTGTTGAAATCGTTGCACAAAGGCGCTCTTGAGCTTATCAAG 454
Qy 121 AlaArgAlaAlaThrHisAlaArgTrpLysSerLeuGluGluArgValAlaHisHis 140
Db 455 GTCACGACAGCTTACCAAGCTCGATACAGAAATCATCGAGAAATGTCGGCAACAC 514
Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTrpArgTrpGluGly 160
Db 515 ACATCTGGTGACCTTGTGAAGCTCTTGCTTCCCTCTTGAGACCTTTCAGATGAGGA 574
Qy 161 AspGluValAsnMetThrLeuAlaLysGluAlaLysLeuValHisGluLysIleLys 180
Db 575 GATGATGTGACATGATGCTTGCAAGATCTGAAGCTAAGATCTTACGACGAAAGTCTCA 634
Qy 181 AspLysHisThrSerAsnArgLysAspValIleArgTrpLeuSerThrArgSerLysAlaGlu 200
Db 635 GAGAAATCTTACAGTACGATGCGATGCTTCAATCAGAAATCTTGACAGAAAGCAAGACAG 694
Qy 201 IleAspAlaThrPheAsnArgTrpGluAspAspHisGlyGluGluLysLeuLysSerLeu 220
Db 695 CTCGGTGACCACTCAACCCACTCAACAGAGATGAGAAACCCATTAAACAGACTTG 754
Qy 221 GluGluGluLysAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGluCysLeu 240
Db 755 AAGGAAGAGTCGACGACCAATGACTCATGAAACTGTAAGACCTTAATCATCATGTGTTG 814
Qy 241 ThrArgProGluLeuLysTrpPheValAspValLeuArgSerAlaIleAsnLysThrArgLys 260
Db 815 ACATTCCTCGAAGAGCAATTTTGAAAGGTTCTTGCTATCATCAACAAATGGGAAACA 874
Qy 261 AspGluGlyAlaLeuThrArgIleValIleThrThrArgAlaGluIleAspLeuLysValIle 280
Db 875 GAGGATGGGGAGTAAACCCAGTCGTGACTACACGAACTGAGATGATGAAACCCATC 934
Qy 281 GlyGluGluTrpGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspTrp 300
Db 935 AAGAGGAAATATCAGCGAAGAAACACACATTCCTTGGACCGTGTCTATCCGCAAGACACT 994
Qy 301 ArgGlyAspTrpGluLysMetLeuValAlaLeuLeuGlyGluAspAspAla 317
Db 995 TCTGTGACTATGAGACATGCTTGTGCTTCTCGACATGGCGATGCT 1045

RESULT 15
AAC40038
ID AAC40038 standard; DNA; 1160 BP.
XX
XX AAC40038;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 26820.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
```

```
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX 25-FEB-1999; 99US-0121825P.
XX 25-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0125785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132486P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134370P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137223P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0138119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140692P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
```

```
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-015088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147312P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 22-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.

PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 3.07e-98
Score: 1067.00
Percent Similarity: 79.81%
Best Local Similarity: 64.04%
Query Match: 66.23%
DB: 3
Gaps: 0

US-10-690-564-2 (1-317) x AAC40038 (1-1160)

QY 1 MetAlaThrLeuIyValSerAspSerValProAlaProSerAspAlaGluGlnLeu 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 96 ATGGCGCTCTCAAAAGTCCCAAGCAGCATGTTCCCTTCCCGAAGATACCCGCAACTC 155

QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 156 CACAAAGCTTTTTCAGAGATGGGTACCAAGAAAGACTATATCAATACCTACCTCAC 215

QY 41 ArgSerAlaGluGlnArgIyValIleArgIlnAlaTyrHisGluThrTyrGlyIuAsp 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 216 AGGAACGACGACCAAGCAGCAGCTTGATCCGAGGTTATGACGACTACATGAGAT 275

QY 61 LeuLeuIyThrLeuAspIySglIleuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 276 CTCTCAAGCATTTACAAAGAGCTTTTACGAGACTTGTAGAGAGCTGTGATGTGTGG 335

QY 81 ThrLeuGluProGlyIuArgAspAlaLeuLeuAlaAsnGluAlaThrIyAspArgTrpThr 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 336 ACTCTTGATCCACCAAGAGAGATGCTTATTGGCTTAAGAAATCCACCAAGATGTTACAC 395

QY 101 SerSerAsnGlnValIleuMetGluValAlaIaCyThrArgThrSerThrGlnLeuLeuHis 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 396 AAGAACAAATGGGTTCTTGTGAATCGCTTGACACAAAGCCGTGCTTACGCTTATCAAG 455
```

```
Qy 121 AlaArgGlnIleAlaThrHisAlaArgTyrLysSerLeuGlnIleAspValAlaHisHis 140
Db 456 GTCAAGCAGCTTACCAAGCTCGATTACAGAAATCAATCGAGAAAGTGTGGCGCAAC 515
Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlnGly 160
Db 516 ACATCTGGTACCTTGTAGCTCTTGCTTCTCTGTGTAGCACTTCAGGTATGAAGGA 575
Qy 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGlnLysIleLys 180
Db 576 GATGATGTGAACATGATGCTGCAAGCTGAAGCTAAAGTACTTCACGAAGGCTCTCA 635
Qy 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Db 636 GAGAAATCCATCAAGTACGATGACTTCAATCAGAAATCTTGACAAACAGAAAGCAACAG 695
Qy 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlnGluLysIleLeuLysSerLeu 220
Db 696 CTCGCTGCACACTCAACCACTACAAACAGAGATGGAACGCCATTAAACAAGAACTTG 755
Qy 221 GlnGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db 756 AAGGAAAGTCGAGACGAACTACATGAAACTACTAAGAGCTGTAATCACATGTTTG 815
Qy 241 ThrArgProGlnLeuTyrPheValAspValIleuArgSerAlaIleAsnLysThrGlyThr 260
Db 816 ACATACCTCTAGAAAGCATTTTGAGAGGTTCTTCGTCTATCAATCAACAAATGGGACA 875
Qy 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGlnIleAspLeuLysValIle 280
Db 876 GACGAATGGGGACTAACCCGAGTCGTGACTACACGAAGTGAAGTTGACATGGAACGCATC 935
Qy 281 GlnGluGlnLysArgArgArgSerIleProLeuGlnLysAlaIleThrLysAspThr 300
Db 936 AAAGAGGAATATCAGCAAGAAACAGCATTCCTTTGAGCCGCTCTATCGCCAAAGACACT 995
Qy 301 ArgGlyAspTyrGlnLysMetLeuValAlaLeuLeuGlnGlyGluAspAspAla 317
Db 996 TCTGGTACTATGAGCAATGCTTGTGCTCTTCTCGGACATGGCGATGCT 1046
```

Search completed! August 22, 2005, 10:30:32  
Job time : 677 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 07:26:26 ; Search time 3736 Seconds

(without alignments)  
3229.757 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611  
Sequence: 1 MATLKXSDSPAPSDDAEQL.....KOTRGDYKMKVALLGEDDA 317

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-Q/cgml2.1/USPTO.spool/US10690564/runat.19082005.172307.25304/app.query.fasta\_1.455  
-DB=EST -QPMF=fastap -SUFFIX=rc -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10690564\_@CGML2.1.1.3437\_@runat.19082005.172307.25304 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: +  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1611	100.0	1095	3	CNSOABR2 BX816579 Arabidops
2	1579	98.0	1137	3	CNSOABR2 BX815936 Arabidops
3	1561	96.9	1155	3	CNSOABR2 BX818120 Arabidops
4	1520	94.4	1034	3	CNSOABR2 BX81810 Arabidops
5	1502	93.2	910	7	CF652924 CF652924
6	1489	92.4	1063	5	CF652924 CF652924
7	1406	87.3	843	5	CF652048 CF652048
8	1329	82.5	814	6	CA782049 CA782049
9	1327	82.4	966	3	CNSOABR3 BX815939 Arabidops

10	1257.5	78.1	782	5	BU635010	BU635010 003A04 In
11	1228.5	76.3	725	5	BU636350	BU636350 049H05 In
12	1188.5	73.8	929	6	CB350421	CB350421 P1F02 Col
13	1137.5	70.6	889	7	CO121757	CO121757 GR_EB03B
14	1111.5	69.0	896	7	CO129429	CO129429 GR_EB28E
15	1109.5	68.9	881	7	CO116938	CO116938 GR_EB019
16	1103.5	68.5	855	7	CO116932	CO116932 GR_EB019
17	1095	68.0	660	6	CB256182	CB256182 89-EB01274
18	1084.5	67.3	861	7	CO128796	CO128796 GR_EB25N
19	1084.5	67.3	863	7	CO128013	CO128013 GR_EB12N
20	1083.5	67.3	870	7	CO132053	CO132053 GR_EB44K
21	1082.5	67.2	880	7	CO116065	CO116065 GR_EB018
22	1080.5	67.1	868	7	CO132772	CO132772 GR_EB45M
23	1078.5	66.9	864	7	CO128843	CO128843 GR_EB25O
24	1077.5	66.9	866	7	CO124740	CO124740 GR_EB07N
25	1075	66.7	672	5	BP561140	BP561140 BP561140
26	1067.5	66.3	933	7	CK266662	CK266662 EST712740
27	1067	66.2	1107	3	CNSOABR1	BX829954 Arabidops
28	1064.5	66.1	867	6	CB893541	CB893541 EST646333
29	1056.5	65.6	979	7	CK277220	CK277220 EST723298
30	1048.5	65.1	841	7	CO113247	CO113247 GR_EB013
31	1048.5	65.1	849	7	CO121318	CO121318 GR_EB02H
32	1046.5	65.0	843	7	CO121288	CO121288 GR_EB02G
33	1046.5	65.0	850	4	BM359402	BM359402 GA_Ba001
34	1043.5	64.8	847	6	CB894292	CB894292 EST647084
35	1043.5	64.8	901	7	CO128478	CO128478 GR_EB25E
36	1039.5	64.5	841	7	CO073880	CO073880 GR_Ea31L
37	1037	64.4	865	3	CNSOABR1	BX814350 Arabidops
38	1036.5	64.3	880	7	CO117199	CO117199 GR_EB019
39	1029.5	63.9	903	7	CO074048	CO074048 GR_Ea33P
40	1027.5	63.8	938	7	CV271171	CV271171 WS0153_B2
41	1027	63.7	661	5	BP560951	BP560951 BP560951
42	1020.5	63.3	818	7	CO131729	CO131729 GR_EB44C
43	1015	62.9	878	7	CNSOABR2	BX829955 Arabidops
44	1013.5	62.9	915	7	CK263099	CK263099 EST709177
45	1013.5	62.9	915	7	CK267919	CK267919 EST713997

#### ALIGNMENTS

RESULT 1  
CNSOABR2 1095 bp mRNA linear HTC 06-FRB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLTPGH56ZG07 of Hormone Treated Callus of strain col-0 of  
Arabidopsis thaliana (thale cress).

ACCESSION BX816579.1 GI:42471684  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidops.

1 (bases 1 to 1095)  
Castell, V., Aury, J.M., Jallion, O., Wincker, P., Clapet, C.,  
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,  
Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished  
2 (bases 1 to 1095)

REFERENCE  
Genoscope.

Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

The sequences are based on single pass reads.  
Full-length libraries (a division of Invitrogen) members carried out

life-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castell, V., Aury, J.M., Jallion, O., Wincker, P., Menard, M., Cruaud, C.,

Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information Center for Protein  
 Sequences) . 5 prime and 3 prime are assembled with Phrap.  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full1](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full1)  
 length  
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

## FEATURES

source

Location/Qualifiers  
 1..1095  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLTRGH56ZG07"  
 /tissue\_type="Hormone Treated Callus"  
 /plasmid="PCMVSPORT\_6"  
 1..1095  
 /gene="At1g35720"

## gene

## ORIGIN

## Alignment Scores:

Pred. No.:	2e-180	Length:	1095
Score:	1611.00	Matches:	317
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-10-690-564-2 (1-317) x CNSOABR2 (1-1095)

```

QY 1 MetAlaThrLeuLySValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20
   |||||
Db 23 ATGGCACTTTAAAGTTTCTGATTCTGTCCTGCTCTTCTATGATGCTGACAAATTG 82
   |||||
QY 21 ArgThrAlaPheGluGlyTyrGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
   |||||
Db 83 AGAACCGCTTTTGAAGATGGGGTACGACGACGAGACTGTATCATCAATCTTGGCTCAC 142
   |||||
QY 41 ArgSerAlaGluGluAlaGlyValIleArgGluAlaTyrHisGluThrTyrGlyGluAsp 60
   |||||
Db 143 AGAAGTGTCTAAACAGAGAAAGTCATCAGCAAGCACTACACAGAACTACCGCGAAGAC 202
   |||||
QY 61 LeuLeuLyThrLeuAspLyGluLeuSerAspPheGluAlaIleLeuLeuTyr 80
   |||||
Db 203 CTTCTCAAGACTCTTGACAGAGACTCTTAACGATTTCCAGAGAGCTATCTTGTGTGG 262
   |||||
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLySArgTyrThr 100
   |||||
Db 263 ACTCTGAACCCGGTGAAGCGTGAATGCTTTATGCTATGACCTCAAAAAGATGAGACT 322
   |||||
QY 101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGluLeuLeuHis 120
   |||||
Db 323 TCAGCAACCAAGTCTTATGAGAGTCTTTCACAGACGATCAACGCGCTGCTTCC 382
   |||||
QY 121 AlaArgGluAlaTyrHisAlaArgTyrLySerSerLeuGluGluAspValAlaHisHis 140
   |||||
Db 383 GCTAGGCAAGCTTACCATGCTCCCTACAGAAAGTCTTTAAAGAGAGCTTGTCTCAC 442
   |||||
QY 141 ThrThrGlyAspPheArgLySLeuLeuValSerLeuValThrSerTyrArgTyrGlyGly 160
   |||||
Db 443 ACTACCGGTACTTCAGAAAGCTTTGTTCTTCTTGTACTCATPACAGTACGAGGA 502
   |||||
QY 161 AspGluValAsnMetThrLeuAlaLySLeuGluAlaLySLeuValHisGluLySLeuHis 180
   |||||
Db 503 GATGAAGTGAACATGACATTGGCTTAACAGAAAGCTTAAGCTGTGTCATGAGAAATCAAG 562
   |||||
QY 181 AspLySHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLySAlaGlu 200
   |||||
Db 563 GACAGACACTACATGATGAGATGATTTATTAATCTTGTCCACAGAGCAAAAGCTCAG 622
   |||||
QY 201 IleAsnAlaThrPheAsnArgTyrGluAspAspHisGlyGluGluIleLeuLySLeu 220
   |||||

```

```

Db 623 ATCAATGCTACTTTTAACGTTACCAAGATGATCAGCGAGAAATTTCTCAAGACTCTT 682
   |||||
QY 221 GluGluGlyAspAspAspAspLySLeuAlaLeuLeuAspSerThrIleGluGlyLeu 240
   |||||
Db 683 GAGGAGAGGAGATGATGACAAAGTTCCTTGACCTTTGAGGTCAACCATTCAGTGTGG 742
   |||||
QY 241 ThrArgProGluLeuTyrPheValAspValIleuArgSerAlaIleAsnLySThrGlyThr 260
   |||||
Db 743 ACAGACCGAGACTTTACTTCTTGTCCATGCTTCTTCTTACGACATCAACAAACTGGA 802
   |||||
QY 261 AspGluGlyAlaLeuThrArgIleValAlaThrThrArgAlaGluIleAspLeuLySValIle 280
   |||||
Db 803 GATGAAGGAGCACTCACTGATGATTTGACCAAGAGCTGAGATTGACTTGAAGCTCAAT 862
   |||||
QY 281 GlyGluGlyTyrGluThrArgArgAsnSerIleProLeuGluGlyValIleThrLySAspThr 300
   |||||
Db 863 GGAAGAGATACCAAGCGACGAAACGACATCTTTGAGAAAGCTATTATCCAAAGACACT 922
   |||||
QY 301 ArgGlyAspTyrGlyLeuMetLeuValAlaLeuLeuGlyGluAspAspAla 317
   |||||
Db 923 CGTGAGATTACGAGAAAGATGCTCGTCCGACTTTCGGTGAAGATGATGCT 973
   |||||

```

RESULT 2  
 CNSOABR2 1137 bp mRNA linear HTC 06-FEB-2004  
 LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 DEFINITION GSLTRGH18ZF02 of Hormone Treated Callus of strain col-0 of  
 Arabidopsis thaliana (thale cress).

ACCESSION BX815936  
 VERSION BX815936.1 GI:42472184  
 KEYWORDS HTC; GS/LT cDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana

REFERENCE Castelli V., Aury J.M., Jallion O., Wincker P., Clepet C., Schachter V., Menard M., Cruaud C., Querier F., Scarpelli C., Salanoubat M., Temple G., Caboche M., Weissenbach J. and Salanoubat M.  
 Whole Genome Sequence Comparisons and Full-Length cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome  
 Annotation

## TITLE

JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1137)  
 AUTHORS Direct Submission  
 TITLE Submitted (18-NOV-2003) Genoscope - Centre National de Séquençage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 JOURNAL - Web : www.genoscope.cns.fr)

## COMMENT

The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelli  
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information Center for Protein  
 Sequences) . 5 prime and 3 prime are assembled with Phrap.  
[http://www.genoscope.cns.fr/externe/sequences/Banque\\_Projet\\_EF/Full1](http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full1)  
 length  
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

## FEATURES

source

Location/Qualifiers  
 1..1137  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSLTRGH18ZF02"  
 /tissue\_type="Hormone Treated Callus"  
 /plasmid="PCMVSPORT\_6"  
 1..1137  
 gene

ORIGIN /gene="At1g35720"

## Alignment Scores:

Pred. No.: 1,34e-176 Length: 1137  
 Score: 1579.00 Matches: 313  
 Percent Similarity: 96.62% Conservative: 1  
 Best Local Similarity: 96.31% Mismatches: 1  
 Query Match: 98.01% Indels: 10  
 DB: 3 Gaps: 1

US-10-690-564-2 (1-317) x CNS0ABT1 (1-1137)

```

Qy 3 ThrlenuyValSerAsp-----SerValProAla 12
   |||||
Db 4 ACCTAAAGATAGAGAAATATGAAAGACTCTTAAAGATTAATAAAATCTGTTCCGCT 63
   |||||
Qy 13 ProSerAspAAlaGluGlnLeuArgThrAlaPheGluGlyTrpGlyThrAsnGluAsp 32
   |||||
Db 64 CTTTCTGATGATGCTGAGCAATTGAGAACCGCTTTTGAAGATGGGCTACGAACGAGAC 123
   |||||
Qy 33 LeuIleSerIleLeuAlaHisArgSerAlaGluGlnArgIleValIleArgGlnAla 52
   |||||
Db 124 TTGATCATATCAATCTTGCTCAAGAAAGTGTGAACAGAGAAAGTCATCAGGCAAGCA 183
   |||||
Qy 53 TyrHisGluThrTyrGlyGluAspLeuLeuLeuThrLeuAspGlyGluLeuSerAspAsp 72
   |||||
Db 184 TACCAAGAAACCTAGCGGAGAACCTTCTCAAGACTTGAACAAGAGCTCTCTTAACGAT 243
   |||||
Qy 73 PheGluArgAlaIleLeuLeuLeuTrpThrLeuGluProGlyGluArgAspAlaLeuLeuAla 92
   |||||
Db 244 TTGAGAGAGACTATCTTGTGTGACCTTGAACCCGCTAGCTGATGCTTTATTTGGCT 303
   |||||
Qy 93 AsnGluAlaThrIleValArgTrpThrSerSerAsnGlnValLeuMetGluValAlaCysThr 112
   |||||
Db 304 AATGAAGCTACAAAGATGAGACTTCAAGCAACCAAGATTCTTATGAAAGTTGCTTGACA 363
   |||||
Qy 113 ArgThrSerThrGlnLeuLeuHisAlaArgGlnAlaTyrHisAlaArgTyrIleValSer 132
   |||||
Db 364 AGAACAATCAACGACGCTGCTTCAAGCTTGAAGCAAGCTTCAATGCTTCAAGAAAGTCT 423
   |||||
Qy 133 LeuGluGluAspValAlaHisIleThrThrGlyAspPheArgIleLeuLeuValSerLeu 152
   |||||
Db 424 CTTGAAGAGACGTTGCTCACACACTACCGGTGACTTCCGAAGACTTTGGTTCTCTT 483
   |||||
Qy 153 ValThrSerTyrArgTyrGluGlyAspGluValAsnMetThrLeuAlaIleArgGlnAla 172
   |||||
Db 484 GTTACCTCATACAGGTACGAAGGAGATGAATGACATGACATTTGGCTTAAAGAAAGCT 543
   |||||
Qy 173 LysLeuValHisGluValIleValAspIleHisTyrAsnAspGluAspValIleArgIle 192
   |||||
Db 544 AACCTGGTTCATAGAAATCAAGAGCAACACCTACATATATGAGATGTTATTAGATC 603
   |||||
Qy 193 LeuSerThrArgSerIleValGlnIleAsnAlaThrPheAsnArgTyrGlnAspAspHis 212
   |||||
Db 604 TTGTCCACAGAAAGCAAGCTCAGATCAATGCTTAAACGTTTACCAAGATGATCAT 663
   |||||
Qy 213 GlyGluGluIleLeuValSerLeuGluGluGlyAspAspAspAspIlePheLeuAlaLeu 232
   |||||
Db 664 GCGGAGGAATTTCTCAAGAGCTTTGAGAGAGAGATGATGATGATCAAGATTCCTTGACCT 723
   |||||
Qy 233 LeuArgSerThrIleGlnCysLeuThrArgProGluLeuTyrPheValAspValLeuArg 252
   |||||
Db 724 TTAGGTCAACCAATTCAGTCTTGAACAAGCAAGAGCTTATCTTGTGATGATGTTCCGT 783
   |||||
Qy 253 SerAlaIleAsnValThrGlyThrAspGluGlyValAlaLeuThrArgIleValIleThrArg 272
   |||||
Db 784 TCAGCAATCAACAAACTGGAATGATGAAGGACCACTCTGAATTTGACACACAAAGA 843
   |||||
Qy 273 AlaGluIleAspLeuValIleGlyGluGluTyrGlnArgArgAsnSerIleProLeu 292
   |||||
Db 844 GCTGAGATGATGTAAGGTATCATTTGAGAGAGATGACAGGCGCAAGAAACATTCCTTTG 903

```

Qy 293 GluLysAlaIleThrIleAspThrArgGlyAspTyrGlyIleMetLeuValAlaLeuLeu 312  
 |||||  
 Db 904 GAGAAAGCTATTTCCAAAGACACTCGTGAAGATTACGAAAGATGCTCGTCGACTTCTC 963

Qy 313 GlyLysAspAspAla 317  
 |||||  
 Db 964 GGTGAAGATGATGCT 978

## RESULT 3

CNS0AB51 1155 bp mRNA linear HTC 06-FEB-2004  
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 GS1R162ZB06 of Silique of strain col-0 of Arabidopsis thaliana  
 (thale cress).

ACCESSION EX818120.1 GI:42471185  
 VERSION EX818120.1  
 KEYWORDS HTC; GS1T cDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana

REFERENCE  
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 1155)

TITLE  
 AUTHORS Menard,M., Crnaud,C., Queller,F., Scarpelli,C., Schachter,V.,  
 Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.,  
 Whole Genome Sequence Comparisons and 'Full'-length cDNA Sequences:  
 A Combined Approach to Evaluate and Improve Arabidopsis Genome  
 Annotation

JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1155)  
 AUTHORS Genoscope.  
 TITLES Direct Submission  
 JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
 BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr

COMMENT  
 The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Castelii  
 V., Aury J.M., Jallion O., Wincker P., Menard M., Crnaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information Center for Protein  
 Sequences) . 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
 length  
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

## FEATURES

source  
 1..1155  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GS1R162ZB06"  
 /issue\_type="Silique"  
 /plasmid="pCMVSPORT\_6"  
 1..1155  
 /gene="At1g35720"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.9e-174 Length: 1155  
 Score: 1561.00 Matches: 310  
 Percent Similarity: 98.43% Conservative: 3  
 Best Local Similarity: 97.48% Mismatches: 4  
 Query Match: 96.90% Indels: 1  
 DB: 3 Gaps: 0

US-10-690-564-2 (1-317) x CNS0AB51 (1-1155)

Qy 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAAlaGluGlnLeu 20



```

Db      14 ATGGCGACTCTTAAGTTTCGATTCTTCGCTCCTCTGATGATGCTGAGCAATTG 73
Qy      21 ArgThrAlaPheGluGluYrTrpGlyThrAangluAspLeuIleIleSerIleLeuAlaHis 40
Db      74 AGAACCCGCTTTTGAAGATGGGGTACCAAGAGAGACTGTATCATCAATCAATCTGGCTCAC 133
Qy      41 ArgSerAlaGluGluGluGluValIleArgGlnAlaIleArgIleGluThrTrpGlyGluAsp 60
Db      134 AGAAGTTCTGAACGAGGAAGATCATCAGCCAAAGCATACCAAGAACCTACGGCGGAAGAC 193
Qy      61 LeuLeuYsThrLeuAspIlysgIuLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80
Db      194 CTTCTCAAGACTCTTGACAAAGAGCTCTTAACGATTTCCAGAGAACTATCTGTTGGTGG 253
Qy      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIlyAspTrpThr 100
Db      254 ACTCTTGAACCCGGTGGCGTTATGCTTTATGGCTTATATGAACTCAAAAAGATGACT 313
Qy      101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db      314 TCAGCAACCAAGTTCTTATGAAAGTTGCTTGCAACAGACATCAACGACGCTGCTTAC 373
Qy      121 AlaArgGlnAlaIleArgHisAlaIleArgIlyLeuYsSerLeuGluAspValAlaHisHis 140
Db      374 GCTAGGCAAGCTTACCATGCTCGCTACAGAAAGTTTCTGAAGAGAGCGTTGCTCACAC 433
Qy      141 ThrThrGlyAspPheArgIlyLeuLeuValSerLeuValThrSerIlyArgTrpGlyGly 160
Db      434 ACTACCGGTACTTCAAAAGCTTTGGTTTCTTGTATCCATCAACAGTTCAGAAAGA 493
Qy      161 AspGluValAsnMetThrLeuAlaIlyGlnGluAlaIlyLeuValHisGluYsIleYs 180
Db      494 GATGAAGTGAACATGATGATGGCTTAAGCAAGAGAGCTAGCTGATGAGAAATCAAG 553
Qy      181 AspIlyHisIlyTrpAsnAspGluAspValIleArgIleLeuSerThrArgSerIlyAlaGln 200
Db      554 GACAGACACTTACATGATGATGAGATGTTATGAAATCTTGTCCACAAAGAACAAAGCTCAG 613
Qy      201 ILeAsnAlaThrPheAsnArgTrpGlnAspAspHisGlyGluGluIleLeuYsSerLeu 220
Db      614 ATCAATGCTACTTTTAAACCGTTTACCAAGATGATCATGCGAGAAATTTCCAAAGACTCT 673
Qy      221 GlnGluGlyAspAspAspAspIlyPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db      674 GAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733
Qy      241 ThrArgProGluLeuTrpPheValAspValLeuArgSerAlaIleAsnIlyThrGlyThr 260
Db      734 ACAAGACCAAGAGCTTTACTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 793
Qy      261 AspGluGlyAlaLeuThrArgGlyIleValThrThrArgAlaGluIleAspLeuYsVal-11 280
Db      794 GATGAAGAGACACTCATGATTAATTTGACCAAGAGACTAAATTTAACTTGAAGGCTCCAT 853
Qy      280 eglYgluGluTrpGlnArgArgAsnSerIleProLeuGluYsAlaIleThrIlyAspThr 300
Db      854 TGAAGAGAGATACCAAGGACGAGAACAGCATTCCTTTGAGAAAGCTATTTCCAAAAAAC 913
Qy      300 rArgGlyAspTrpGlyIlySerMetLeuValAlaLeuLeuGlyGluAspAspAla 317
Db      914 TCGTGAGATTAACGAGAAATGCTCTCGACATTTCTCGTGAGAGATGATGATGAT 965

```

```

RESULT 4
CNS0913M      1034 bp      mRNA      linear      HTC 06-FEB-2004
LOCUS      Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION      GSLTUS23B01 of Adult vegetative tissue of strain col-0 of
ACCESSION      Arabidopsis thaliana (thale cress).
VERSION      BX841810.1 GI:42454468
KEYWORDS      HTC; GSLT_cDNA.
SOURCE      Arabidopsis thaliana (thale cress)

```

```

ORGANISM      Arabidopsis thaliana
REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE          rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsids.
JOURNAL        1 (bases 1 to 1034)
JOURNAL        Castellani V., Aury J.M., Jaillon O., Wincker P., Clepet C.,
JOURNAL        Menard M., Cruaud C., Queller F., Scarpelli C., Schachter V.,
JOURNAL        Temple G., Caboche M., Weissenbach J., and Salanoubat M.
JOURNAL        Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
JOURNAL        A Combined Approach to Evaluate and Improve Arabidopsis Genome
JOURNAL        Annotation
JOURNAL        Unpublished
JOURNAL        2 (bases 1 to 1034)
JOURNAL        Genoscope.
JOURNAL        Direct Submission
JOURNAL        Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
JOURNAL        BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
JOURNAL        - Web : www.genoscope.cns.fr)
COMMENT        The sequences are based on single pass reads.
COMMENT        Life Technologies (a division of Invitrogen) members carried out
COMMENT        full-length libraries construction : Temple G.
COMMENT        Genoscope members carried out sequencing and annotation : Castellani
COMMENT        V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
COMMENT        Schachter V., Weissenbach J., Salanoubat M.
COMMENT        URGV INRA : Clepet C., Caboche M.
COMMENT        Annotation is based on the June 2003 version of the Arabidopsis
COMMENT        genome released by MIPS (Munich Information center for Protein
COMMENT        Sequences). 5 prime and 3 prime are assembled with Phrap.
COMMENT        http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
COMMENT        length
FEATURES
source          location/Qualifiers
1..1034
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTUS23B01"
/tissue_type="adult vegetative tissue"
/plasmid="pCMVSPORT_6"
1..1034
/gene="Atlg35720"

```

```

ORIGIN
Alignment Scores:
Pred. No.:      1 21e-169      Length:      1034
Score:          1520.00      Matches:      302
Percent Similarity: 98.41%      Conservative: 7
Best Local Similarity: 96.18%      Mismatches: 5
Query Match:    94.35%      Indels:      1
DB:              3      Gaps:      0

```

```

US-10-690-564-2 (1-317) x CNS0913M (1-1034)
Qy      1 MetAlaThrLeuIlyValSerAspSerValProAlaProSerAspAlaGluGlnLeu 20
Db      23 ATGGGAGACTTTAAGTTTCTGATTCGTTCTGCTCGCTTGAAGATGCTGAGCAATTG 82
Qy      21 ArgThrAlaPheGluGluYrTrpGlyThrAangluAspLeuIleIleSerIleLeuAlaHis 40
Db      83 AGAACCCGCTTTTGAAGATGGGGTACCAAGAGAGACTGTATCATCAATCAATCTGGCTCAC 142
Qy      41 ArgSerAlaGluGluGluGluValIleArgGlnAlaIleArgIleGluThrTrpGlyGluAsp 60
Db      143 AGAAGTCTGAACAGAGAAAGTCAATCAGGCAAGCATACCAAGAGCTTACGCGGAAGAC 202
Qy      61 LeuLeuYsThrLeuAspIlysgIuLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80
Db      203 CTTCTCAAGACTCTTGACAAAGAGAGCTCTTAACGATTTTGAAGAGAGCTATCTTGGTTGG 262
Qy      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIlyAspTrpThr 100
Db      263 ACTCTTGAACCCGGTGGCGGTGATGCTTTATGGCTTATATGAGAGCTTACAAAGAGATGACT 322

```

Qy 101 SerSerAenGlnValIleuMerGluValAlaCyThrAArgThrSerSerThrGlnIleuLeuHis 120  
 Db 323 TCAAGCAACCAAGTTCTTATGGAGATTCTTGCAAGAGACATCAAGCGAGCTGCTTCAC 382  
 Qy 121 AlAaRGlnAlaTyRHISAlAaRGTyLySylSerLeuGluAaPValAlaHis 140  
 Db 383 GCTAGGCAAGCTTACCATGCTCGCTCAAGAGCTCTTGAGAGAGAGCTTGCTCACAC 442  
 Qy 141 ThrThrGlyAaPheAaGlyLeuLeuValSerLeuValThrSerTyRArgTyRGlulGly 160  
 Db 443 ACTACCGGTAAGCTTCAAGAGAGCTTTGGTTCTCTTGTAACCTCATACAGTACGAAGA 502  
 Qy 161 AaPGLValAaMetThrLeuAlaLyGlnGluAlaLySylLeuValHisGluLySylLeu 180  
 Db 503 GATGAAGTGAACATGACATGGCTTACCAAGAACCTAAGCGGTCCTAGAGAGATCAG 562  
 Qy 181 AaPlyHisTyRaaAaPGLAaPValAlAaRGTyLeuSerThrAaSerTyRAlaGln 200  
 Db 563 GACAGACATCAATCATGATGAGATGTTATTAGATCTTGCCACAGAGAGAGCTCAG 622  
 Qy 201 IlAaSnAlaThrPheAaAaRGTyRGlAaAaPHisGlyGluGluIleuLySylLeu 220  
 Db 623 ATCAATCTCACTTTGGCCGCTTACCAAGATGATCATGCGAGGGAATTCACAGAGCTT 682  
 Qy 221 GluGluGluLyAaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 240  
 Db 683 GAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 742  
 Qy 241 ThrAaPProGluLeuTyRPhaValAaPValLeuAaPValLeuAaPValLeuAaPValLeu 260  
 Db 743 ACAAGACCAAGCTTCTTCTTGCTGATGATGATGATGATGATGATGATGATGATGATG 802  
 Qy 261 AaPGLuGlyAlaLeuThrAaRGlyLeuValThrThrAaGlnIleAaPLeuLySylLeu 280  
 Db 803 GATGAGGAGACACTCACTAAGATTGTGACCAAGAGACTGAGATTGATGAGAGTCAAT 862  
 Qy 281 GlyGluGluLyTyRGlAaAaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 300  
 Db 863 GAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 922  
 Qy 301 ArgGlyAaPlyTyRGlulLySylMetLeuValAlaLeuLeuGluGlu 314  
 Db 923 CTTGAGATTAACAAGAGAG-CTGCTGCACTACTCATGAG 963

RESULT 5  
 CF652924 910 bp mRNA linear EST 06-NOV-2003  
 LOCUS 85-1020135W-066-001-121-SP6P MP12-ADIS-066 Arabidopsis thaliana  
 DEFINITION CDNA clone MP12p200112110 5'-PRIME, mRNA sequence.  
 ACCESSION CF652924  
 VERSION CF652924.1 GI:37429878  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana

REFERENCE  
 AUTHORS Schmidt,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,  
 Mitchell-Olds,T. and Weisshaar,B.  
 TITLE large-scale identification and analysis of genome-wide  
 JOURNAL single-nucleotide polymorphisms for mapping in Arabidopsis thaliana  
 MEDLINE Genome Res. 13 (6), 1250-1257 (2003)  
 PUBMED 22683290  
 COMMENT Contact: Weisshaar B  
 ADIS DNA core facility at MP12  
 Max-Planck-Institute for Plant Breeding Research  
 Carl-von-Linne Weg 10, 50829 Koeln, Germany  
 Fax: 00492215062851  
 Email: weisshaar@mpiz-koeln.mpg.de  
 Insert Length: 910 Std Error: 0.00

Plate: 1 row: 1 column: 21  
 Seq primer: SP6P.  
 Location/Qualifiers  
 1. 910  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Ws-0"  
 /db\_xref="GABI:939529"  
 /db\_xref="taxon:3702"  
 /clone="MP12p200112110"  
 /issue\_type="root"  
 /lab\_host="E. coli TOP10"  
 /clone\_lib="MP12-ADIS-066"  
 /note="vector: pCMVSPORT6; Site\_1: SalI; Site\_2: NotI;  
 cDNA library from Arabidopsis thaliana, accession  
 W881lewk1ja-0; roots from three weeks old plants grown  
 on MS-plates at 26M-0C with 16 hours light/day; library  
 was made at the Max-Planck-Institute for Plant Breeding  
 Research, Cologne, Germany; cloning sites SalI-NotI,  
 primer sites and orientation:  
 SP6-SalI-CCACGCGTCCG-5prime-cDNA-polyA-CC-NotI-T7; GATEWAY  
 compatible; Note: Sequencing granted in the context of the  
 GABI Arabidopsis Verbund I: Genetic Diversity.  
 'Establishment of high-efficiency SNP-based mapping tools  
 and development of methods for genome-wide mutation  
 detection' PI: Bernd Weisshaar Sequence submission managed  
 by RZPD/GABI-Primary database: <http://gabi.rzpd.de> This  
 clone is available from RZPD; contact RZPD ([clone@rzpd.de](mailto:clone@rzpd.de))  
 for further information."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,4e-167 Length: 910  
 Score: 1502.00 Matches: 296  
 Percent Similarity: 99.66% Conservative: 0  
 Best Local Similarity: 99.66% Mismatches: 1  
 Query Match: 93.23% Indels: 0  
 DB: 7 Gaps: 0

US-10-690-564-2 (1-317) x CF652924 (1-910)

Qy 1 MetAlaThrLeuLyValSerAaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 20  
 Db 18 ATGGCCACTTTTGAAGGTTCTGATTTCTGTCCTGCTCTTGATGATGATGCAATGG 77  
 Qy 21 ArgThrAlaPheGluGlyTyRGlTyRaaGlnAaPLeuIleSerIleLeuAlaHis 40  
 Db 78 AGAACCGCTTTTGAAGAGATGGGATGAGAGAGACTTATCATCATCTTGCGTCAAC 137  
 Qy 41 ArgSerAlaGluGlnAaRGlyValIleAaRGlnAlaTyRHISGluThrTyRGlulAaP 60  
 Db 138 AGAAGTGTGAACAGAGAAAGTCAATCAGGCAAGCATCAATGAACCTACGCGCAAGAC 197  
 Qy 61 LeuLeuLyThrLeuAaPlySylLeuSerAaAaPPhaGlnAaGlnAlaIleLeuLeuTrp 80  
 Db 198 CTTTCAACACTTTTGAACAGAGAGCTCTTCAACACTTGAAGAGACTATCTTTGTTGG 257  
 Qy 81 ThrLeuGluPProGlyGluAaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaaPaa 100  
 Db 258 ACTCTTGAACCCGGTGAAGGTATGCTTATTGCTTATGAGAGAGAGAGAGAGAGTACT 317  
 Qy 101 SerSerAenGlnValIleuMerGluValAlaCyThrAaRGTyRThrAaSerThrGlnLeuHis 120  
 Db 318 TCAAGCAACCAAGCTTATGGAGATTCTTGCAAGAGACATCAAGCGAGCTGCTTCAC 377  
 Qy 121 AlAaRGlnAlaTyRHISAlAaRGTyLySylSerLeuGluAaPValAlaHisHis 140  
 Db 378 GCTAGGCAAGCTTACCATGCTCGCTCAAGAGCTCTTGAGAGAGAGCTTGCTCACAC 437  
 Qy 141 ThrThrGlyAaPheAaGlyLeuLeuValSerLeuValThrSerTyRArgTyRGlulGly 160  
 Db 438 ACTACCGGTAAGCTTCAAGAGAGCTTTGGTTCTCTTGTAACCTCATACAGTACGAAGA 497

QY 161 ASPGLVAlaAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
| | | | |  
| | | | |  
Db 498 GATGAAGTGAACATGACATTGGCTAAGCAAGAAAGCTAGTGCATGAGAAATCAAG 557  
| | | | |  
QY 181 ASPLyHisrTyraAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200  
| | | | |  
| | | | |  
Db 558 GACACAGACATCAATCATGTAGTAGATGTTATATGAAATCTTGCACCAAGAACCAAGCTCAG 617  
| | | | |  
QY 201 ILeAsnAlaThrPheAsnArgTyrgLInAspAspHisGlyGlnGluIleLeuLysSerLeu 220  
| | | | |  
| | | | |  
Db 618 ATCAATGCTACTTTTAAACGCTTACCAAGATGATCAGCGAAGAAATTTCCAAAGACTTT 677  
| | | | |  
QY 221 GlnGlnGlyAspAspAspAspLysPheLeuAlaLeuLysSerThrIleGlnCysLeu 240  
| | | | |  
| | | | |  
Db 678 GAGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 737  
| | | | |  
QY 241 ThrArgProGluLeuTyrgPheValAspValLeuLysSerAlaIleAsnLysThrGlyThr 260  
| | | | |  
| | | | |  
Db 738 ACAAGACCAAGAGCTTTAATTGTCGATGTTCTTCTGTCAGCAATCAACAAACTGGAACCT 797  
| | | | |  
QY 261 ASPGlnGlyAlaLeuThrArgIleValIleThrThrArgAlaGluIleAspLeuLysValIle 280  
| | | | |  
| | | | |  
Db 798 GATGAAGAGACCTCTCACTAATATTGTCACCAAGAGCTGAGATTGACTTGAAGCTCAT 857  
| | | | |  
QY 281 GlyGlnGluTyrgLInArgArgAsnSerIleProLeuGlnLysAlaIleThr 297  
| | | | |  
| | | | |  
Db 858 GGAAGAGAGTACCAGCGCAGAAACAGCATTTCTTTGGAGAAAGCCATTACC 908  
| | | | |

RESULT 6  
BX838972 1063 bp mRNA linear EST 11-FEB-2004  
LOCUS BX838972 Arabidopsis thaliana flowers and buds Col-0 Arabidopsis  
DEFINITION thaliana cDNA clone GSLTFRB502D03 SPRIM, mRNA sequence.  
VERSION BX838972  
KEYWORDS BX838972.1 GI:42533055  
SOURCE EST.  
ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana  
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

## REFERENCE

AUTHORS

## TITLE

A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
Unpublished (2004)

## JOURNAL

COMMENT

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jailion O., Wincker P., Menard M., Cnauud C.,  
Schachter V., Weisenbach J., Salanoubat M.  
JRGV INRA : Clepet C., Caboche M.  
Genoscope is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences).  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/EST  
http://www.genoscope.cns.fr/cgi-bin/g9b/g9b?source=Arabidopsis.

## FEATURES

Source

1. 1063  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLTFRB502D03"  
/tissue\_type="flowers and buds"

## ORIGIN

/clone\_11b="Arabidopsis thaliana flowers and buds Col-0"

## Alignment Scores:

Pred. No.:	6,11e-166	Length:	1063
Score:	1489.00	Matches:	292
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	92.43%	Indels:	0
DB:	5	Gaps:	0

US-10-690-564-2 (1-317) x BX838972 (1-1063)

QY 26 GlnTyrgLysThrAsnGluAspLeuIleIleSerIleLeuAlaHisArgSerAlaGln 45  
| | | | |  
| | | | |  
Db 2 GGATGGGGTACGACAGAGAGACTTGATCATATCAATCTTGGCTCAGAGAAAGTCTGACAG 61  
| | | | |  
QY 46 ArgLysValIleArgGlnAlaTyrgLysGluValAspLeuLysThrLeu 65  
| | | | |  
| | | | |  
Db 62 AGGAAAGTCATCAGGCAAGCATACCAAGAACTACGGGGAAGACCTTCTCAAGACTCTT 121  
| | | | |  
QY 66 AspLysGluLeuSerAspAspPheGluArgAlaIleLeuThrThrLeuGluProGly 85  
| | | | |  
| | | | |  
Db 122 GACAAAGAGCTCTTACGATTTCCAGAGAGCTATCTTGTGTGACTTTGAAACCGGAT 181  
| | | | |  
QY 86 GluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrgLysSerSerAsnGlnVal 105  
| | | | |  
| | | | |  
Db 182 GACCGTGATGCTTTATTTGGCTTAATGAGCTCAAAAGATGAGCTTCAAGCAACCAAGT 241  
| | | | |  
QY 106 LeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHisAlaArgGlnAlaTyrg 125  
| | | | |  
| | | | |  
Db 242 CTTAATGAGAGTGGCTTGCACAAAGACATCAAGAGAGCTGCTTCAAGGAGCAAGCTTAC 301  
| | | | |  
QY 126 HisAlaArgTyrgLysSerLeuGlnGluAspValAlaHisHisThrThrGlyAspPhe 145  
| | | | |  
| | | | |  
Db 302 CAGGCTCGCTACAAAGAGTCTCTTGAAGAGAGCTGCTCACACACCTACCGGTGCTTC 361  
| | | | |  
QY 146 ArgLysLeuLeuValSerLeuValThrSerTyrgTyrgLysGlnGlyAspGluValAsnMet 165  
| | | | |  
| | | | |  
Db 362 AGAAAGCTTTTGGCTTCTCTTCTTACCTCATACAGTTCAGAAAGAGATGAAGCAAG 421  
| | | | |  
QY 166 ThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLysAspLysHisrTyraAsn 185  
| | | | |  
| | | | |  
Db 422 ACATTGGCTTAACCAAGAAAGCTAGCTGCTCATGAGAAATCAAGAGACATCAAT 481  
| | | | |  
QY 186 AspGluAspValIleArgIleLeuSerThrArgSerLysAlaGlnIleAsnAlaThrPhe 205  
| | | | |  
| | | | |  
Db 482 GATGAGGATGTTATTGAATCTTGTCCACAAAGAGCAAGCTCAGATCAATCTACTTT 541  
| | | | |  
QY 206 AsnArgTyrgLInAspAspHisGlyGlnGluIleLeuLysSerLeuGlnGlyAspAsp 225  
| | | | |  
| | | | |  
Db 542 AACCGTTACCAAGATGATGATGCGAGGAAATTTCCAAAGAGCTTGAAGAAAGAGATGAT 601  
| | | | |  
QY 226 AspAspLysPheLeuAlaLeuLysSerThrIleGlnCysLeuThrArgProGluLeu 245  
| | | | |  
| | | | |  
Db 602 GATGCAAAATTCCTTGCACATTTTGAAGCTCAACCAATTCAGTGTCTTCAAGAACAGAGCTT 661  
| | | | |  
QY 246 TyrgPheValAspValLeuLysSerAlaIleAsnLysThrGlyThrAspGlnGluValLeu 265  
| | | | |  
| | | | |  
Db 662 TACTTGTGATGATCTTCTGTTCCACAAATCAAAAGCTGGAAGCTGATGAAGAGAGACTC 721  
| | | | |  
QY 266 ThrArgIleValThrThrArgAlaGluIleAspLeuLysValIleGlyGlnGluTyrgLIn 285  
| | | | |  
| | | | |  
Db 722 ACTAAGATTTGATCCACAAAGAGCTGATGATGATGATGATGATGATGATGATGATGATG 781  
| | | | |  
QY 286 ArgArgAsnSerIleProLeuGlnLysValaIleThrLysAspThrArgGlyAspTyrgLIn 305  
| | | | |  
| | | | |  
Db 782 CCAGAGAACAGATTCCTTGGAGAAAGCTATTACCAAGACATCTGTGAGATTTACGAG 841  
| | | | |  
QY 306 LysMetLeuValAlaLeuLeuLysGluAspAspAla 317  
| | | | |  
| | | | |  
Db 842 AAGATGCTGTGACACTTCTCGTGAAAGATGATGCT 877  
| | | | |

RESULT 7  
CF652048  
LOCUS  
DEFINITION  
30-L020380-066-004-L08-SP6P MP1Z-ADIS-066 Arabidopsis thaliana cDNA  
CF652048  
VERSION  
CF652048.1 GI:37428171  
KEYWORDS  
SOURCE  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.  
1 (bases 1 to 843)  
Schmid, K.J., Sørensen, T.R., Stracke, R., Torjek, O., Altmann, T.,  
Mitchell-Olds, J., and Weisshaar, B.  
TITLE  
Large-scale identification and analysis of genome-wide  
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana  
Genome Res. 13 (6), 1250-1257 (2003)  
JOURNAL  
MEDLINE  
22683290  
PUBMED  
12799357  
COMMENT  
Contact: Weisshaar B  
ADIS DNA core facility at MP1Z  
Max-Planck-Institute for Plant Breeding Research  
Carl-von-Linne Weg 10, 50829 Köln, Germany  
Email: weisshaar@mpiz-koeln.mpg.de  
Insert Length: 843 Std Error: 0.00  
Plate: 4 row: L column: 08  
Seq primer: SP6P.  
Location/Qualifiers  
1..843  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Ws-0"  
/db\_xref="GABI:938637"  
/db\_xref="taxon:3702"  
/clone="MP1Zp2001L084Q"  
/cissue\_type="root"  
/lab\_host="E. coli TOP10"  
/clone\_1db="MP1Z-ADIS-066"  
/note="Vector: PCMVSPORT6; Site 1: Sall; Site 2: NotI;  
cDNA library from Arabidopsis thaliana, accession  
Masliewekija-0; roots from three weeks old plants grown  
on WS-plants at 26M-0C with 16 hours light/day; library  
was made at the Max-Planck-Institute for Plant Breeding  
Research, Cologne, Germany; cloning sites Sall-NotI,  
primer sites and orientation:  
SP6-Sall-CCACGCGCCG-5prime-cDNA-polyA-CC-NotI-T7; GATEWAY  
compatible; Note: Sequencing granted in the context of the  
GABI Arabidopsis Verbund II: Genetic Diversity,  
'Establishment of high-efficiency SNP-based mapping tools  
and development of methods for genome-wide mutation  
detection'. PI: Bernd Weisshaar Sequence submission managed  
by RZPD/GABI-Primary database: <http://gabi.rzpd.de> This  
clone is available from RZPD; contact RZPD (clone@rzpd.de)  
for further information."

ORIGIN  
Alignment Scores:  
Pred. No.: 3,24e-156 Length: 843  
Score: 1406.00 Matches: 276  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 87.27% Indels: 0  
DB: 7 Gaps: 0  
US-10-690-564-2 (1-317) x CF652048 (1-843)

QY 1 MetAlaThrLeuIyValSerAspSerValProAlaProSerAspAlaGluInLeu 20  
Db 14 ATGGCGACCTTAAGGTTCTGATTCTGTTCTGCTCTTCTGATGATGCTGAGCAATTG 73

QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40  
Db 74 AGAACCGCTTTTGAAGATGGGTACCAACGAGACTTGATCATCATCTTGGCTCAC 133  
QY 41 ArgSerAlaGluGlnArgValIleArgGlnAlaTryHisGluThrTryGlyGluAsp 60  
Db 134 AGAAGTGTGAACAGAGGAAGATCATCAGGCAAGCATACATGAACTTCGCGCAAGAC 193  
QY 61 LeuLeuIyThrLeuAspIyGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTrp 80  
Db 194 CTTCCTCAAGCTCTTACCAAGAGGCTCTTACAGACTTCGAGAGACTATCTTGTTGTGG 253  
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIyAspGlyTrp 100  
Db 254 ACTCTGAAACCCGTGACGTGATGCTTATTGCTTAAAGACTTCGAGAGACTATCTTGTTGTGG 313  
QY 101 SerSerAsnGluValLeuMetGluValAlaCyThrArgThrSerThrGlnLeuLeuHis 120  
Db 314 TCAGCAACCAAGTGTATGGAAGTTGCTTGCAAGAGCATCAACGACGCTTCAC 373  
QY 121 AlaArgGlnAlaTryHisAlaArgTylIyIySerLeuGluGluAspValAlaHisHis 140  
Db 374 GCTAAGCAAGCTTACCATGCTGCTTCAAGAACTCTTGAAAGAGACGTTGCTCAC 433  
QY 141 ThrThrGlyAspPheArgIySerLeuLeuValSerLeuValThrSerTryArgTylGly 160  
Db 434 ACTACCGGATCTTCGAAAGCTTTGGTTCTCTTGTTACCTCATACAGTGCAGAGGA 493  
QY 161 AspGluValAsnMetThrLeuAlaIyGlnGluAlaIySerValIleGluIyIyLeys 180  
Db 494 GATGAAGTGAACATGACATTTGGTTAAGCAAGAGCTTACGTGTCATGAAATCAAG 553  
QY 181 AspIyHisIyTryAsnAspGluAspValIleArgIleLeuSerThrArgSerIyAlaGln 200  
Db 554 GACAAACCTACATCATGATGATGATGATTAGATCTTGTCCCAAGAAAGCAAGCTCAG 613  
QY 201 IleAsnAlaThrPheAsnArgTylIyIyAspAspHisGlyGluGluIleuIySerLeu 220  
Db 614 ATTAATGCTTACTTTAACGTTTACCAAGATGATCAGGAGAAATTTCTCAAGAGCTTT 673  
QY 221 GluGluGluIyAspAspAspAspIyPheLeuAlaLeuLeuArgSerThrIleGlnIyLeu 240  
Db 674 GAGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 733  
QY 241 ThrArgProGluLeuIyTryPheValAspValIleuArgSerAlaIleAsnIyThrGlyTrp 260  
Db 734 ACAAGACCAAGAGCTTACTTGTTCGATGTTCTTGTTCGACGATCAACAAACTGGAAC 793  
QY 261 AspGluGluIyAlaLeuThrArgIleValThrThrArgAlaGluIleAsp 276  
Db 794 GATGAAGAGCACTCACTGAATGTCACCAAGAGCTGAGATTGAC 841

RESULT 8  
CA782049  
LOCUS  
DEFINITION  
01B11AR Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA  
CA782049 814 bp mRNA linear EST 03-DEC-2002  
VERSION  
CA782049.1 GI:26020095  
KEYWORDS  
SOURCE  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsie.  
1 (bases 1 to 814)  
Lundgaard, M., Emmersen, J., Nielsen, K.L., Wilson, I., Somerville, S.,  
and Weindler, K.G.  
TITLE  
EST sequencing of Erysiphe cichoracearum infected Arabidopsis  
plants  
JOURNAL  
Unpublished (2002)  
COMMENT  
Institute for biotechnology

Aalborg Universitet  
 Sohngaardsholmsvej 49, 9000 Aalborg, Denmark  
 Tel: +45 96358467  
 Fax: +45 98141808  
 Email: kym@bio.au.dk

## FEATURES

source

Location/Qualifiers  
 1. .814  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Columbia"  
 /db\_xref="taxon:3702"  
 /dev\_stage="Plant 3 weeks old, three days post infection"  
 /clone\_lib="Infected Arabidopsis leaf"  
 /note="Organ: Leaf; Vector: pBluescript; Mixed cDNA library of Arabidopsis and E. chloracanthum infected leaf from three weeks old Arabidopsis plants. Plants were harvested 3 days after infection and mRNA oligo dt selected."

## ORIGIN

## Alignment Scores:

Pred. No.:	4,39e-147	Length:	814
Score:	1329.00	Matches:	267
Percent Similarity:	98.89%	Conservative:	0
Best Local Similarity:	98.89%	Mismatches:	2
Query Match:	82.50%	Indels:	2
DB:	6	Gaps:	0

US-10-690-564-2 (1-317) x CA782049 (1-814)

QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20  
 7 ATGGCGACTTGTAG-GTTCTGATTCCTGCTTCCTTCGATGATGTCGACATGG 65  
 QY 21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuLeileSerlleLeuAlaHis 40  
 66 AGAACCGCTTTTAAAGATGGGTAGCAGACGACTGATCATCAATCTTGGCTCAC 125  
 QY 41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrlHisGlnThrTyrlGlyGluAsp 60  
 126 AGAAGTGTCTAACAAGAGAGATCATCAGGCAAGCATCCACAGAACTTCGGCGAAGAC 185  
 QY 61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluAlaIleLeuLeuTyr 80  
 186 CTTCTCAAGACTCTTTCAGACGAGCTCTTCAACGATTCGAGAGACTATCTTGTGTGG 245  
 QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTTPThr 100  
 246 ACTCTTGAACCCGGTGGCGTGTATGCTTATTTGGCTAATGAAGCTACGAAAGATGACT 305  
 QY 101 SerSerAsnGlnAlaLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
 306 TCAGACACCAAGTTCTTATGAGAGTTGCTTGCACAGACATCAACGACGCTGCTTCAAC 365  
 QY 121 AlaArgGlnAlaTyrlHisAlaArgTyrlLysSerLeuGluGluAspValAlaHisHis 140  
 366 GCTAGACAGACTTACATGATGCTCGCTACAGAAAGCTCTTGAAGAGAGCGTGTCTCAAC 425  
 QY 141 ThrThrGluAspPheArgLysLeuLeuValSerLeuValThrSerTyrlArgTyrlGlyGly 160  
 426 ACTACCGGTACTTCGAAAGCTTTGGTTCCTTCTGTTAACCCTCAACAGGTACGAGGA 485  
 QY 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
 486 GATGAAGTGAACATGACATTGGCTTAAGCAAGAAAGCTTAAGTGTCCATGAGAAATCAAG 545  
 QY 181 AspLysHisTyrlAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200  
 546 GACAAGACATCAATGATGAGATGTTATTAGATCTTGTCCACAGAGCAAGCAAGCTCAG 605  
 QY 201 TLeAsnAlaThrPheAsnArgTyrlGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220  
 606 ATCAATGCTACTTTTAAACGCTTACCAAGATGATCGCGAGGAAATTTCTCAAGAGCTTT 665

QY 221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
 666 GAGGAAGAGATGATGATGATCAACAAGTCTCTGACATTTGAGGTCAACCATTTAGTCTTG 725  
 QY 241 ThrArgProGluLeuTyrlPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260  
 726 ACAAGACCAAGACTTACTTGTCTGATGTTCTTCTGTCGATCATCAACAAACTGGAAC 785  
 QY 261 -AspGluGlyAlaLeuThrArgIleVal 269  
 786 GCGTGAAGGAGCATCATCAAGATTGTG 813  
 Db 786 GCGTGAAGGAGCATCATCAAGATTGTG 813

RESULT 9  
 CDS0ABV3  
 LOCUS  
 DEFINITION  
 CDS0ABV3 966 bp mRNA linear HTC 06-FEB-2004  
 Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
 GSTRPGH182H06 of Hormone treated Callus of strain col-0 of  
 Arabidopsis thaliana (thale cress).

ACCESSION BX815939.1 GI:42472187

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

COMMENT

Unpublished  
 2 (bases 1 to 966)  
 Direct Submission  
 Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
 BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 The sequences are based on single pass reads.  
 Life Technologies (a division of Invitrogen) members carried out  
 full-length libraries construction : Temple G.  
 Genoscope members carried out sequencing and annotation : Caetelli  
 V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,  
 Schachter V., Weissenbach J., Salanoubat M.  
 URGV INRA : Clepet C., Caboche M.  
 Annotation is based on the June 2003 version of the Arabidopsis  
 genome released by MIPS (Munich Information center for Protein  
 Sequences). 5 prime and 3 prime are assembled with Phrap.  
 http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
 length  
 http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

## FEATURES

source

1. .966  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Col-0"  
 /db\_xref="taxon:3702"  
 /clone="GSTRPGH182H06"  
 /tissue\_type="Hormone treated Callus"  
 /plasmid="pCMVSPORT\_6"  
 1. .966  
 /gene="AT1G35720"

## ORIGIN

## Alignment Scores:

Pred. No.:	9,68e-147	Length:	966
Score:	1327.00	Matches:	261
Percent Similarity:	99.62%	Conservative:	1
Best Local Similarity:	99.24%	Mismatches:	1
Query Match:	82.37%	Indels:	0

DB:	3	Gap:	0
US-10-690-564-2 (1-317) x CNS0ABV3 (1-966)			
OY	55	GIUThTyrXG1LGIuAspLeuLeuYbThrLeuAspLySG1uLeuSerAspAspHeG1u	74
Db	2	GAGACCTACGGCGAAGACCTTCTCAAGACTCTTGACAGAGACTCTTAACGATTTCCAG	61
OY	75	ArgAlaIleLeuLeuTrpThrLeuG1uProG1uAArgAspAlaLeuLeuAlaAsnG1u	94
Db	62	AGAGCTACTTGTGTGGACTCTTGAAACCCGGTGAAGCGTGAAGCTTTATTGGCTAATGA	121
OY	95	AlaThrIlybAArgTrpThrSerSerAsnG1uValLeuMetG1uValAlaCysThrArgTrh	114
Db	122	GCTACAAAAGATGAGCTTCAAGCAACCAAGTTCTTAAGAAAGTTGCTTGACAAAGACA	181
OY	115	SerThrG1nLeuLeuHAlaAlaArg1nAlaTyrH1sa1aArgTyrLybSerLeuG1u	134
Db	182	TCAAGCGACGCTGCTTCACGCTAGGCAAGCTTACCAATGCTGCTACAAAGCTCTTGA	241
OY	135	GIuAspValAlaAlaH1eH1eThrThrG1yAspPheArgLybLeuLeuValSerLeuValThr	154
Db	242	GAGAGCGTTCCTCACACACACATACCGGTGACTTCAGAAAAGCTTTTGCTTCTTACC	301
OY	155	SerTyrArgTyrG1uG1yAspG1uValAsnMetThrLeuAlaLybG1nAlaLybLeu	174
Db	302	TCAATACAGTACGAAGAGATGAATGAATGAATGACATTGGCTAAGCAAGAGCTAAGCTG	361
OY	175	ValH1eSG1uLyb11eLybAspLybH1eTyrTrAsnAspG1uAspVal11eArg11eLeuSer	194
Db	362	GTCACATGGAAGAAATCMAAGACAAAGACATCAATGATGATGATGATTAATGATCTTGTCC	421
OY	195	ThrArgSerLybAlaG1n11eAsnAlaThrPheAsnArgTyrG1nAspAspH1eG1yG1u	214
Db	422	ACAAGAACCAAGCTCAAGTCAATCTACTTTTAACCGTTACCAAGATGATCAATGGCAG	481
OY	215	GIu11eLeuLybSerLeuG1uG1uG1yAspAspAspLybPheLeuAlaLeuLeuArg	234
Db	482	GAAATTTCCCAAGAGCTTGAAGGAAGATGATGATGATCAAGTTCTTGACCTTTTGGG	541
OY	235	SerThr11eG1nCybLeuThrArgProG1uLeuTyrPheValAspValLeuArgSerAla	254
Db	542	TCAACCATTCAGTGTCTTCAAGAACCAAGAGCTTACTTTGTCCGAAGTCTTCGTTCAGCA	601
OY	255	11eAsnLybThrG1yThrAspG1uG1yAlaLeuThrArg11eValThrThrArgAlaG1u	274
Db	602	ATCAACAAACCTGGAACCTAGTGAAGAGACCTCACTAATGATGATGATCAACAAAGCTGAG	661
OY	275	11eAspLeuLybVal11eG1yG1uG1uTyrG1nArgArgAsnSer11eProLeuG1uLyb	294
Db	662	ATTGACTTGAAAGGTACTATTGGAAGGAGTACCAAGCCAGCATTCCTTTGGAGAA	721
OY	295	Ala11eThrIlybAspTrhArgG1yAspTyrG1uLybMetLeuValAlaLeuLeuG1yG1u	314
Db	722	GCTATTACCAAGAACTCGTGGAGATTACGAGAAGAGTCTCGTGCACCTTCGGTGAA	781
OY	315	AspAspAla 317	
Db	782	GATGATGCT 790	
RESULT 10			
LOCUS	BUE35010	782 bp	mRNA linear EST 23-SEP-2002
DEFINITION	003A04 Infected Arabidopsis leaf	Arabidopsis thaliana cDNA, mRNA	
ACCESSION	BUE35010		
VERSION	BUE35010.1	GI:23302265	
KEYWORDS	EST.		
SOURCE	Arabidopsis thaliana (chale crese)		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.		

REFERENCE	1 (bases 1 to 782)
AUTHORS	Lundsgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S. and Welinder,K.G.
TITLE	EST sequencing of Erysiphe cichoracearum infected Arabidopsis plants
JOURNAL	Unpublished (2002)
COMMENT	Contact: Karen G. Welinder Institut for bioteknologi Aalborg Universitet Sohnegaardsholmvej 49, 9000 Aalborg, Denmark Tel: +45 96358467 Fax: +45 98141808 Email: Kgw@bio.auc.dk.
FEATURES	Location/Qualifiers
source	1..782
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/ecotype="Columbia"
	/db_xref="taxon:3702"
	/dev_stage="Plant 3 weeks old, three days post infection"
	/clone_lib="Infected Arabidopsis leaf"
	/note="Organ: Leaf; Vector: pBluescript; Mixed cDNA library of Arabidopsis and E. cichoracearum infected leaves from three weeks old Arabidopsis plants. Plants were harvested 3 days after infection and mRNA oligo dT selected."
ORIGIN	
Alignment Scores:	
Pred. No.:	1,31e-138
Score:	1257.50
Percent Similarity:	99.22%
Best local Similarity:	99.22%
Query Match:	78.06%
DG:	5
	Gaps: 0
US-10-690-564-2 (1-317) x BU635010 (1-782)	
OY	1 MetAlaThrLeuIyValSerAspSerValProAlaProSerAspAspAlaGlunLnu 20
DB	19 ATGGCGACTTTAAAGT-TCTGAATTCGT-CCTGCCTCTTGAGATGCTGACCAATTG 76
OY	21 ArgThrAlaPheGluGlyTrpGlyThraAngLuAspLeuIleIleSerIleLeuAlHis 40
DB	77 AGAACCGCTTTGAAGAAGTGGGTGCGAACGAGCATTTGATCATAATCATCTTGCTCAC 136
OY	41 ArgSerAlaGluInlmarGlyValIleArgInalatyRhIsGIuThrTYRGLyGuAsp 60
DB	137 AGAAGTGTGACAGAGAAAGTCTACAGGACAGCATGCCAGAACTCCAGCGGAAAGAC 196
OY	61 LeuLeuLyThrLeuAspLyseGluLeuSerAsnAspPheGluwaGuaAlaIleLeuLeuTrp 80
DB	197 CTTCACAAGCTTTGACAAAGAGCTCTTCAAGATTTGAGAGAGAGCTATCTTGTTGG 256
OY	81 ThrLeuGluProGlyGuIyurAspAlaLeuLeuAlaEnguAlaThrLyAsxTrpThr 100
DB	257 ACTCTTGAACCCGGTAGCGGTATGCTTTATGGCTAATGAAGCTACAAAAGATGACT 316
OY	101 SerSerAsnGlnValLeuMetGluValAlaCySerTharGthrSerThrGlnLeuLeuHis 120
DB	317 TCAGGCAACCAAGTTCTTATGGAAGTGTCTTGACAAAGACATCAACGACGCTCTTACC 376
OY	121 AlaArgGlnAlaTyRhIaAlaArgTyRlybysSerLeuGluGluwArpValAlaHisHis 140
DB	377 GTTAGGCAAGCTTACATGCTTGACAAAGAGCTCTTGAAAGAGAGCTTGCTGACACAC 436
OY	141 ThrThGlyASPPhetqLyseLeuValSerLeuValThrseryTrarqTyrgLugly 160
DB	437 ACtAcCGGTGACTTCGAAGAGCTTTGGTTTTCTCTGTACTCTATKACAGTACGAAGA 496
OY	161 AspGluValAsnMetThrLeuAlaLyGlnGluValaLyseLeuValHisGluLyVilLeLy 180
DB	497 GAAGAAGTAATGACATGGCTTGAAGCAAAAGCTGAGTGGTCCATGAGAAATCAAG 556



Qy 181 AsplyhiSTYrAsnaSpGluAspValIleArgIleuSerThrArgSerLySAIagIn 200  
 Db 557 GACAAGACACTACAAATGATGAGATGTTATTAAGATCTTGCCACAGAACCAAGCTCAG 616  
 Qy 201 ILeaenAlaThrPheAanaArgTYrGlnAspAspHisGlyGluGluIleLeuLySserLeu 220  
 Db 617 ATCAATGCTACTTCTTAAACCGTTACCAAGATGATCATGGCAGAGAAATTCACAGAGTCTT 676  
 Qy 221 GluGluGluAspAspAspAspAspAspAspAspAspAspAspAspAspAspAspAspAsp 240  
 Db 677 GAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 736  
 Qy 241 ThrArgProGluLeuLeuTYrPheValAspValIleuArgSerLySAIaIle 255  
 Db 737 ACAAGACCGAGCTTACTTGTGTGAGTGTCTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 781

RESULT 11  
 BU636350 775 bp mRNA linear EST 23-SEP-2002  
 LOCUS 049H05 Infected Arabidopsis leaf Arabidopsis thaliana cDNA, mRNA  
 DEFINITION  
 sequence.  
 ACCESSION BU636350 GI:23303605  
 VERSION BU636350.1  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE 1 (bases 1 to 775)  
 AUTHORS Lundgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S. and Wellinder,K.G.  
 TITLE EST sequencing of Erysiphe cichoracearum infected Arabidopsis plants  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Karen G. Wellinder  
 Institut for bioteknologi  
 Aalborg Universitet  
 Sohngaardsholmsvej 49, 9000 Aalborg, Denmark  
 Tel: +45 96358467  
 Fax: +45 98141808  
 Email: kgw@bio.auc.dk.

FEATURES  
 source location/Qualifiers  
 1..775  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /ecotype="Columbia"  
 /db\_xref="taxon:3702"  
 /dev\_stage="Plant 3 weeks old, three days post infection"  
 /clone\_lib="Infected Arabidopsis leaf"  
 /note="Organ: Leaf; Vector: pBluescript; Mixed cDNA library of Arabidopsis and E. cichoracearum infected leaf from three weeks old Arabidopsis plants. Plants were harvested 3 days after infection and mRNA oligo dt selected."

ORIGIN  
 Alignment Score: 3.63e-135 Length: 775  
 Pred. No.: 1228.50 Matches: 249  
 Percent Similarity: 98.81% Conservative: 1  
 Best Local Similarity: 98.42% Mismatches: 2  
 Query Match: 76.26% Indels: 3  
 DB: 5 Gaps: 0

US-10-690-564-2 (1-317) x BU636350 (1-775)

Qy 1 MetaIaThrLeuLySValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20  
 Db 19 ATGGCGACCTTTAAAGG-TCTGATTCGT-CTGCTCTTCTGATGATGCTGAGCAATTG 76  
 Qy 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleSerIleLeuAlaHis 40

Db 77 ACAAACCGCTTTGAAGATGGGGTACGAACGAGACTTGATCATATCTTGCTCAC 136  
 Qy 41 ArgSerAlaGluGlnArgValIleArgGlnAlaTYrHisGluThrTrpIleGluAsp 60  
 Db 137 ACAAAGTCTGAACAGAGAAAGTCATCAGGCAAGATACCAAGAACTTCCGCAAGAC 196  
 Qy 61 LeuLeuLYThrLeuAspLySgluSerAsnaAspPheGluArgAlaIleLeuLeuTrp 80  
 Db 197 CTTCCTCAAGACTTTCACAGAGAGCTCTTAACGATTCGAAAGACTATCTGTGTGCG 256  
 Qy 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLySArgTrpThr 100  
 Db 257 ACTCTTGAACCCGGAGAGCGGTATGCTTATTTGGCTTAATGAAGTACAAAGATGAGACT 316  
 Qy 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
 Db 317 TCAAGCAACCAAGTCTTATGGAAGTGTCTGCACAAAGCAATCAACGAGCTGCTTAC 376  
 Qy 121 AlaArgGlnAlaTYrHisAlaArgTYrLySLeuSerLeuGluGluAspValAlaHisHis 140  
 Db 377 GCTAGGCAAGCTTACCATGCTGCTTCACAAAGAGTCTTTGAAGAGAGCTTCTACACAC 436  
 Qy 141 ThrThrGlyAspPheArgLySLeuLeuValSerLeuValThrSerTYrArgTYrGluGly 160  
 Db 437 ACTACCGGTGACTTCAGAAAGCTTTGGTTTCTCTTGTTATCCTCATACAGTACGAAGA 496  
 Qy 161 AspGluValAspMetThrLeuAlaLySngluAlaLySLeuValHisGluLySLeuLyS 180  
 Db 497 GATGAAGTAGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 556  
 Qy 181 AsplyhiSTYrAsnaSpGluAspValIleArgIleLeuSerThrArgSerLySAIagIn 200  
 Db 557 GACAAGACACTACAAATGATGAGATGTTATTAAGATCTTGCCACAGAACCAAGCTCAG 616  
 Qy 201 ILeaenAlaThrPheAanaArgTYrGlnAspAspHisGlyGluGluIleLeuLySserLeu 220  
 Db 617 ATCAATGCTACTTCTTAAACCGTTACCAAGATGATCATGGCAGAGAAATTCACAGAGTCTT 676  
 Qy 221 GluGluGluAspAspAspAspAspAspAspAspAspAspAspAspAspAspAspAspAsp 240  
 Db 677 GAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 736  
 Qy 240 uThrArgProGluLeuLeuTYrPheValAspValIleuArg 252  
 Db 737 GACAAGACCGAGCTTACTTGTGTGAGTGTCTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 773

RESULT 12  
 CB350421 929 bp mRNA linear EST 17-MAY-2003  
 LOCUS J1P02 Cotton fiber subtracted cDNA library Gossypium hirsutum cDNA,  
 DEFINITION  
 mRNA sequence.  
 ACCESSION CB350421 GI:30841101  
 VERSION CB350421.1  
 KEYWORDS EST.  
 SOURCE Gossypium hirsutum (upland cotton)  
 ORGANISM Gossypium hirsutum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
 REFERENCE 1 (bases 1 to 929)  
 AUTHORS Jia,S.J., Luo,Y.C., Feng,J.X., Wei,G., Li,J., Shi,Y.H., Fu,Q., Liu,D., Luo,J.C. and Zhu,Y.X.  
 TITLE Isolation and analyses of genes preferentially expressed during early cotton fiber development by subtractive PCR and cDNA array  
 JOURNAL Nucleic Acids Res. 31 (10), 2534-2543 (2003)  
 MEDLINE 22622070  
 PUBMED 12736302  
 COMMENT Contact: Zhu Y.  
 National Laboratory of Protein Engineering and Plant Genetic Engineering of Life Sciences, Peking University  
 Beijing 100871, China



Tel: 86 10 6275 1193  
Fax: 86 10 6275 4427  
Email: zhuyxewater.pku.edu.cn.  
Location/Qualifiers

# FEATURES

Source

/organism="Gossypium hirsutum"  
/mol\_type="mRNA"  
/db\_xref="taxon:3635"  
/db\_xref="taxon:3635"  
/issue\_type="fiber"  
/dev\_stage="10 days post anthesis (dpa)"  
/clone\_lib="Cotton fiber subtracted cDNA library"  
/note="The library was constructed using PCR-select cDNA subtraction method with 10 dpa cotton fiber as tester and fiberless mutant as driver"

## ORIGIN

### Alignment Scores:

Pred. No.: 2,67e-130 Length: 929  
Score: 1188.50 Matches: 227  
Percent Similarity: 85.76% Conservative: 38  
Best Local Similarity: 73.46% Mismatches: 43  
Query Match: 73.77% Indels: 1  
DB: 6 Gaps: 1

US-10-690-564-2 (1-317) x CB350421 (1-929)

QY 1 MetAlaThrLeuValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20  
DB 4 ATGGCCACTCTTACAGTCCGACGACGACGACGACGACGACGACGACGACGACGACGAC 63  
QY 21 ArgThrAlaPheGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 40  
DB 64 AGAAAGCCTTTTCAGATGGGAACTAATGAGGCTTAATCAATGATATATGGGTGAC 123  
QY 41 ArgSerAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 60  
DB 124 AGAAATCGGAGCAGCAAGAACTTATCGAAAACTTACCGTGAACCTTATGAGAGGAT 183  
QY 61 LeuLeuValThrLeuAspValGluLeuSerAspPheGluGluGluGluGluGluGluGlu 80  
DB 184 CTCTCAAGGACACTGACCAAGAGCTCTCGAATGACTTTGAGAGGCGTTGCTGCTTGG 243  
QY 81 ThrLeuGluProGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 100  
DB 244 GCTCTTATCTCTGTAACGTGATGCCCTTTGGCTATGATGAGCCAAAGGTGACT 303  
QY 101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGluLeuValHis 120  
DB 304 TCAGCAATCAAGTCTTATGAAATGCTGACCAAGTCTGCAACCAAGTCTGCTCAC 363  
QY 121 AlaArgGluAlaValArgValArgValArgValArgValArgValArgValArgValArg 140  
DB 364 GCAAGGAGGCTTATATGCTCGTTATAGAGTCTGTTGAGAGGAGCGTTGCTCATCAC 423  
QY 141 ThrThrGluAspPheAspGluLeuValSerLeuValThrSerValArgValArgValArg 160  
DB 424 ACAGCTGGGAGACTTCCATAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 483  
QY 161 AspGluValAsnMetThrLeuAlaValGluGluGluGluGluGluGluGluGluGluGlu 180  
DB 484 GAGAGGAGTGAACATGACTCTGGCAAAACAGAGCGAGGAGGAGGAGGAGGAGGAGGAG 543  
QY 181 AspValAspValAspValAspValAspValAspValAspValAspValAspValAspVal 200  
DB 544 AACAAAGCTTACAGTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 603  
QY 201 IleAsnAlaThrPheAsnArgValArgValArgValArgValArgValArgValArgVal 220  
DB 604 ATCAATGCAACTTAATCACTCAAAAATGAATATGGAATGACATTAACCAAGGACTTG 663  
QY 221 GluGluGluValAspAspAspAspAspAspAspAspAspAspAspAspAspAspAspAsp 240

DB 664 ---AAGGCTGACCCCTAAGATGAGTTCCTTGACACTAAGTCACAGTGAAGTCTTG 720  
QY 241 ThrArgProGluLeuValPheValAspValLeuValSerAlaValLeuValSerValThr 260  
DB 721 GTCATCCCGGAAAGATTTTGAAGAGTCTCTGCGCTAGCGATCAATACAGAGAACG 780  
QY 261 AspGluGluValLeuValArgValArgValArgValArgValArgValArgValArgVal 280  
DB 781 GATGAAAGAGCTCTTTCAGAGTGTGCTGAGGCTGAGGTGAGTCAATTAAGATCAT 840  
QY 281 GluGluGluValArgValArgValArgValArgValArgValArgValArgValArgVal 300  
DB 841 GCAGATGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
QY 301 ArgGluAspValArgValArgValArgValArgValArgValArgValArgValArgVal 309  
DB 901 CATGAGACTATGAAAAATGCTGCTG 927

## RESULT 13

COL121757 889 bp mRNA linear EST 16-JUN-2004  
LOCUS GR\_EB03B10.f GR\_Eb Gossypium raimondii cDNA clone GR\_EB03B10 5',  
DEFINITION mRNA sequence.  
ACCESSION COL121757 GI:48820444  
VERSION COL121757.1 GI:48820444  
KEYWORDS EST.  
SOURCE Gossypium raimondii  
ORGANISM Gossypium raimondii  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Malvales; Malvaceae; Malvaceae; Gossypium.

REFERENCE 1 (bases 1 to 889)  
Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,  
Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C. and  
Wing, R. A.

TITLE Global assembly of Cotton ESTs  
JOURNAL Unpublished (2004)  
COMMENT Contact: Rod A. Wing  
Arizona Genomics Institute  
The University of Arizona  
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA  
Tel: 520 626 9595  
Fax: 520 621 1259  
Email: http://genome.arizona.edu  
Plate: 03 row: B column: 10.

## FEATURES

source

1. .889  
/organism="Gossypium raimondii"  
/mol\_type="mRNA"  
/db\_xref="taxon:29730"  
/clone="GR\_EB03B10"  
/issue\_type="fiber"  
/dev\_stage="3 to +3 DPA"  
/lab\_host="DH10B"  
/clone\_lib="GR\_Eb"  
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:  
EcoRV; library made by Invitrogen with RNA supplied by  
Wendle lab. Directional cloned into NotI-EV. Clones  
plated/picked by AGI. More glycerol clones held in -80."

## ORIGIN

### Alignment Scores:

Pred. No.: 2.9e-124 Length: 889  
Score: 1137.50 Matches: 218  
Percent Similarity: 86.30% Conservative: 34  
Best Local Similarity: 74.66% Mismatches: 39  
Query Match: 70.61% Indels: 1  
DB: 7 Gaps: 1

US-10-690-564-2 (1-317) x COL121757 (1-889)

QY 1 MetAlaThrLeuValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20

```

Db      14 ATGCCACCTCTACAGTCCCGACGACAGTTCCTTCAAGTGTGAGATGTGAAACAATA 73
Qy      21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
Db      74 AGAAAGCCCTTTTCAGATGGGAACTAATGAGGGCTTAATCATATATATATGGGTGAC 133
Qy      41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaIleArgLysIleuThrTyrGlyGluAsp 60
Db      134 AGAAATCCGAGCAGCAAGAACTTGATTGCAAAAACCTACGCTGAACCTATGAGAGAGAT 193
Qy      61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTyr 80
Db      194 CTCCTCAAGGCACTAGACAGAGAGCTCTCCAAATGACATTTAGAGGGCTGCTTGGCTTGG 253
Qy      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTTPThr 100
Db      254 GCTCTTGATCTCTGTGAACTGATGCCCTTTGGCTTATATAGCCACCAAAAGGTGACT 313
Qy      101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db      314 TCAAGCAATCAGGTCTTATGAAATATGCTGCAAGATCTGCCAACCAACTGCTTAC 373
Qy      121 AlaArgGlnAlaIleTyrHisAlaIleArgTyrLysSerLeuGluAspValAlaHisHis 140
Db      374 GCAAGGAGGCTTATCATGCTCGTTATAGAAGTCGCTGAAGAGATGTTGCTCATCAC 433
Qy      141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160
Db      434 ACGACTGGCCACTTCGTAAGCTTCCTCTACTCTTATGAGTTTCATACATATGAGGGA 493
Qy      161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db      494 GAGAGGTGAACATGACTCTGGCGAAACAGAGCGCAAGTTCCTCATGAGAAATTTCA 553
Qy      181 AspLysHisTyrAsnAspGluAspValIleArgLysLeuSerThrArgSerLysAlaGln 200
Db      554 GACAAACCTTACATGATGACGATGTCATAGGGTTTGGCTACAAAGAACAGAGGACAG 613
Qy      201 ILAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
Db      614 ATCACTCAACTCTGAATCACTACAGAAATGAAATATGAAATGACATAAACAAGACTTG 673
Qy      221 GluGluGluAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Db      674 ---AAGCTGATCTTAAGATGAGTCTCTTGCACATAAGTCCACAGAGAGTGTG 730
Qy      241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
Db      731 GTCTATCCGAAAGATATTTTGAAGAGTCTTCCCTACCAATCATATAGACGAGAACG 790
Qy      261 AspGluGluAlaLeuThrArgLysIleValThrThrArgAlaGluIleAspLeuLysValIle 280
Db      791 GATGAAAGAGCTCTTACTACAGTGTGTTGCACTAGGCTGAGGTTGATCTAAAGATCAT 850
Qy      281 GlyGluGluTyrGlnArgAsnSerIleProLeu 292
Db      851 GCAGATGAGTACCGACGAGGACAGTGTCCCACTG 886

```

```

RESULT 14
LOCUS    COI29429
DEFINITION GR_Eb28E08.r GR_Eb Gossypium raimondii cDNA clone GR_Eb28E08 3',
            mRNA sequence.
ACCESSION COI29429
VERSION    COI29429.1 GI:48878407
KEYWORDS   EST.
SOURCE     Gossypium raimondii
ORGANISM   Gossypium raimondii
            Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Rosidae; eucotyledons; eudicotyledons; core eudicots;
            Rosidae; eucotyledons; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE 1 (bases 1 to 896)
            Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,

```

```

TITLE
JOURNAL
COMMENT
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 28 row: E column: 08.
Location/Qualifiers
FEATURES
source
1..896
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb28E08"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/note="vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; library made by inviterogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Clones
plated/picked by AGI. More glycerol clones held in -80."
ORIGIN
Alignment Scores:
Pred. No.: 3,61e-121 Length: 896
Score: 1111.50 Matches: 215
Percent Similarity: 83.78% Conservative: 33
Best Local Similarity: 72.64% Mismatches: 47
Query Match: 68.99% Indels: 1
DB: 7 Gaps: 1
US-10-690-564-2 (1-317) x COI29429 (1-896)
Qy      1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
Db      9 ATGGCACTCTTACAGTGGCCCGCCAGCCCATTTGGGTGATGTCGAAAGATTGTAAACAATA 68
Qy      21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
Db      69 AGAACTCTTTTTCAGAGTGGGAACTATAGAGGCTTATCATATATATATGGGTGAC 128
Qy      41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaIleArgLysIleuThrTyrGlyGluAsp 60
Db      129 AGAAATGCCGAGCAAGAACTTGATTCGAAAAAACCTACGCTGAACCTATGAGAGAT 188
Qy      61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTyr 80
Db      189 CTCCTCAAGGCACTACAGAGAGCTCTCGAAATGACTTTGAGAGGCTGCTTGGCTTGG 248
Qy      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTTPThr 100
Db      249 GCTCTTGATCTCTGTGAACTGATGCTCTTTGGCTAATGAGCCACCAAAAGTGGACT 308
Qy      101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Db      309 TCAAGCAATCAGGTCTTATGAAATATGCTTCCACAAAGTCTGCCAACCAACTGCTTAC 368
Qy      121 AlaArgGlnAlaIleTyrHisAlaIleArgTyrLysSerLeuGluAspValAlaHisHis 140
Db      369 GCAAGGAGGCTTATCATGCTCTGTTATAGAAGTGCCTTGAAGAGATGTGCTATCAC 428
Qy      141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160
Db      429 ACGACTGGGAGCTTCGTAAGTCTCTCCACTCATAGTATGATCATACAGATATGAGGGA 488
Qy      161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Db      489 GAGGAGGTGAACATGACTCTGGCGAAACAGAGGCGAAGTTCCTTCATAGAAATTTCA 548

```



This Page Blank (uspto)



```
Dh 155 CGGACAGCCTTCGAGGATGGGAGACAAATGAGAGCTGATCATATTCATATTCGGCTCAT 214
Qy 41 ArgSerAlaGluGlnArgIleValIleArgIleAlaTyrHisGlnThrTyrGlyIleuAsp 60
Dh 215 AGGAAATGGCGCGGAGAGAAAGCTGATTCGGCAAACTATGCGCAAGCTTACGGCGAGAGAC 274
Qy 61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTyr 80
Dh 275 CTCCTCAAGGCATTTGGACAGAGAACTTACCAATGATTTGGAGAGCGCGTGGCTTTGG 334
Qy 81 ThrLeuGluProGlyGlyIleArgAspAlaLeuLeuAlaAsnGluAlaThrIlyAspTyrThr 100
Dh 335 TCACCTGATCCGCGCTGAACCTGATGCGTACTCTGGCAATGAAACGAGAAATATGAGACT 394
Qy 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Dh 395 TCAAGCAACAGGATTTTCATGAGAAATAGCCTGACGAGGCTCCGACGCGTGGCTTATG 454
Qy 121 AlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluIleuAspValAlaHisHis 140
Dh 455 GCAAGCAAGCATATGATCCCGATACAGAGATCAATGAGAGAGCGTCCGCTCACAC 514
Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlyGly 160
Dh 515 ACAACTGAGATTTTCTGAAGTTGCTGTAACCTTGGGAGCTCTTACCGTAAATGATGGA 574
Qy 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Dh 575 GATGAGGTGAATGATGATTTGGCAAAAGCAGAGCTTAAGTACTCCAGAGAGAGATCTCA 634
Qy 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Dh 635 GAGAGGCTTATGCGCCATGAGATCTCATAGAGATTTTGGCTACTGAGAGCAAAAGCAC 694
Qy 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
Dh 695 GTCAAATGCTACGCTGATACCTCAAAATGAGTTTGGAAATGATATCAACAGAGATTTG 754
Qy 221 GluGluGluValAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Dh 755 AAAAAT---GATCCAAAAGACGGCTTCTTACTATACAGAGCTACAGTAAAGTCCG 811
Qy 241 ThrArgProGluLeuTyrPheValAspValIleuArgSerAlaIleAsnLysThrArgIlyThr 260
Dh 812 ACTCGCCCTGAGAAAGATTTTGAAGAGTTCTTCGCTACCATCAATTAAGCAGAGAAC 871
Qy 261 AspGluGluValAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValIle 280
Dh 872 GATGAAGGGGCTCTGACCAAGATGATGCTTACACAGGCGCGAGGTTCACATGAAGTTTATA 931
Qy 281 GlyGluGluTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300
Dh 932 AGTGAAGAGTACCAAGAGAGAAATAGCATCTCTGATGCTCCATTTGCAAGAGAACT 991
Qy 301 ArgGlyAspTyrGluLysMetLeuValAlaLeuLeuArgSerThrIleGlnCysLeu 317
Dh 992 ACTGAGAGCATATGAAAAATGCTTCGCGATGATTGGCCAGCTCGAGGCT 1042

RESULT 2
US-09-325-932A-20
; Sequence 20, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
```

```
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-20

Alignment Scores:
Pred. No.: 2,58e-100
Score: 866.50
Percent Similarity: 84.17%
Best Local Similarity: 69.58%
Query Match: 53.79%
DB: 3 Gaps: 1

US-10-690-564-2 (1-317) x US-09-325-932A-20 (1-789)

Qy 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
Dh 73 ATGGCGACTATGCGCGGTGCACCGTCCGTCCGTCCGCGGTGAAGAGATGCCAGCAGCTC 132
Qy 21 ArgThrAlaPheGluGlyTyrGlyIleArgAsnGluAspLeuIleSerIleLeuAlaHis 40
Dh 133 CAAAAAGCTTTGGCAGAGATGGGGGACGAATGAATGATCATATCATACTGCTCAC 192
Qy 41 ArgSerAlaGluGlnArgIleValIleArgGlnAlaTyrHisGluThrTyrGlyIleuAsp 60
Dh 193 AGAAACGACGCGCAGCGAAAGTAAATCCGACAAACATATGCCGAGACATATGGGAGACAT 252
Qy 61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTyr 80
Dh 253 CTTCTCAAGCGCTTGACAGAGAACTCTTACTGACTTTGAGAGATCTTGCTCTGTGG 312
Qy 81 ThrLeuGluProGlyGlyIleArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrThr 100
Dh 313 ACCCTGAGATCCCTGCGGAGGCTGATGCAATCTTGTCCAAATGAAGCTACAGAGATTTGACT 372
Qy 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
Dh 373 TCGAGCAATGGGTTCTCATGAGAAATGCTTGCACAGAGGCTTCAATGAGATTAATTCAG 432
Qy 121 AlaArgGlnAlaTyrHisAlaArgTyrLysLysSerLeuGluGlnAspValAlaHisHis 140
Dh 433 GTGAGCGAGCCTATATCATGCTGTTATGAAGATCTTTGAAGAGACATCCCATATTCAC 492
Qy 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlyGly 160
Dh 493 ACTACTGGGGAATTCGCGAAGCTGCTTGTCTCTGGCAAGTACCTTTCGATAGAGGG 552
Qy 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
Dh 553 CCTGAGGTGAACATGACATTTGGCGAATCAGAGGCTTAAGATCTTCATGAGAAATTCAC 612
Qy 181 AspLysHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
Dh 613 GAGAGGCTTACCAATCATGATGATGCTCATCAAAATTTGTTATCAAGAGATTAAGCTCAG 672
Qy 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGluGluIleLeuLysSerLeu 220
Dh 673 CTTATGCAACCTCATATTAACAAGATGATTTGGAGATGCCATCAACAAGATCTG 732
Qy 221 GluGluGluValAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
Dh 733 ---AAGGCTGATCCAAATGATGAATTTGAAACTGTCGAGATCAGCAATTAAGTGGCTGG 789

RESULT 3
US-09-325-932A-21
; Sequence 21, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE REFERENCE: 1022
```

```
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-21

Alignment Scores:
Pred. No.: 4,43e-63 Length: 704
Score: 573.00 Matches: 109
Percent Similarity: 78.02% Conservative: 33
Best Local Similarity: 59.89% Mismatches: 40
Query Match: 35.57% Indels: 0
DB: 3 Gaps: 0

US-10-690-564-2 (1-317) x US-09-325-932A-21 (1-704)

QY 1 MetAlaThrLeuLySerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
DB 148 ATGTCAACAAATTATAGTCAGTCCATACCGACCCCATCTGAAAGCGCTG 207
QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuLejIleSerIleLeuAlaHis 40
DB 208 AGAAGGCTTTTGAGGGTGAGGCAAAATGAAAGTCATCAATCAATATTAGCAT 267
QY 41 ArgSerAlaGluGlnArgLySerValIleArgGlnAlaIleArgHisGlyThrTyGlyGluAsp 60
DB 268 AGAAGTCAGACACACAGCAAGATATCCGTAAGATTTCACAGCTGCAAGAGCAT 327
QY 61 LeuLeuLyThrLeuAspLySerGluLeuSerAsnAspPheGluArgAlaIleLeuLeuTip 80
DB 328 CTCCTTAGAGCATTAGATCTGAGCTTCAAGTCACTTGAGAAAGCTGTATTCCTTGG 387
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrLySargTrpThr 100
DB 388 GTACTGATCTCAGCTGAACGTGATGCGGTATCTCATGTGCAATAAAGAGTGAAT 447
QY 101 SerSerAsnGlnValIleuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
DB 448 GCAGAGAAATATATCGCTTTAGAAATTTCCAGTGCCTGATCTTCGCTGAACCTATTGATG 507
QY 121 AlaArgGlnAlaIleArgHisAlaArgTyLyLySerLeuGluGluAspValAlaHisHis 140
DB 508 GTGAGGCAAGCATATCATATTCGCTACAAAAGTCCCTCGAAGAGACGTGGCTGCACAT 567
QY 141 ThrThrGlyAspPheArgLySerLeuLeuValSerLeuValThrSerTyArgTyGlyGly 160
DB 568 ACAAGTGAACCTTCGTAAGTGTGGTAGCATTGTAAATTCATATCGGTATGAAGGT 627
QY 161 AspGluValAsnMetThrLeuAlaLySerGlnGluAlaLySerLeuValHisGluLySileLyS 180
DB 628 CCGAAGTGTGATATGATTTGGCAAGTATGAGAGCAAGAAAGCTAAGTGAATCTATTAAC 687
QY 181 AspLyS 182
DB 688 GAGCAA 693

RESULT 4
US-09-325-932A-18
; Sequence 18, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Pflum, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-18

Alignment Scores:
Pred. No.: 1.16e-56 Length: 484
Score: 520.50 Matches: 106
Percent Similarity: 73.29% Conservative: 12
Best Local Similarity: 65.84% Mismatches: 13
Query Match: 32.31% Indels: 31
DB: 3 Gaps: 1

US-10-690-564-2 (1-317) x US-09-325-932A-18 (1-484)

QY 47 LysValIleArgGlnAlaIleArgHisGlyThrTyGlyGluAspLeuLeuLyThrLeuAsp 66
DB 3 AAGCTATTCGGCAACCTATGCGGAGACTTACGGGAGAGACTCTCTCAAGCATTTGAC 62
QY 67 LysGluLeuSerAsnAspPheGluArgAlaIle----- 77
DB 63 AGAAGACTTACCAATGATTTTGAAGT-CTGATCTCTTAATTATTGATTCATCCAT 121
QY 77 ----- 77
DB 122 GGAGACGCGTCCCTCTTCTCAGATTAATCCATATTCATTCCTGATGTCAGAGGCTG 181
QY 78 ---LeuLeuTrpThrLeuGluProGlyGluArgAspAlaLeuLeuAlaHisGlnAlaThr 96
DB 182 GTGGTCTTTGGTGGCTTATCCGCTGAACGATGCGTACTTGGCAATGAAGCGAGC 241
QY 97 LysArgTrpThrSerSerAsnGlnValIleuMetGluValAlaCysThrArgThrSerTrp 116
DB 242 AAAAGTGAACCTTCAGACCAACAGGTTCTCATGAAATAGCCGTCGACAGGCTCCACAG 301
QY 117 GlnLeuLeuHisAlaArgGlnAlaIleArgHisAlaArgTyLyLySerLeuGluGluAsp 136
DB 302 CAGTCTCATGAGCAAGACATATGATGCTCCGATACAGAAAGTCGCTGAAAGAGGAC 361
QY 137 ValAlaHisHisThrThrGlyAspPheArgLySerLeuLeuValSerLeuValThrSerTrp 156
DB 362 GTGCTGCACACACAACTGAAATTTTGTAAAGTGTGCTGATCTTTGTAGCTCTTAC 421
QY 157 ArgTyGluGlyAspGluValAsnMetThrLeuAlaLySerGlnGluAlaLySerLeuValHis 176
DB 422 CATTAATGATGAGATGAGGTGATATGACTTTGGCAAAAGCAGAGGCTAAGATCTCCAC 481
QY 177 Glu 177
DB 482 GAG 484

RESULT 5
US-09-949-016-2081
; Sequence 2081, Application US/0949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2081
```



```

; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2081

Alignment Scores:
Pred. No.:      8.4e-50      Length:      1332
Score:          472.00      Matches:      111
Percent Similarity: 56.86%      Conservative: 63
Best Local Similarity: 36.27%      Mismatches: 130
Query Match:      29.30%      Indels:      2
DB:               4          Gaps:      2

US-10-690-564-2 (1-317) x US-09-949-016-2081 (1-1332)
QY      12  AlAProSerAspAlaGluGlnLeuArgThrAlaPheGluGlyTrpGlyThrAsnGln 31
      99  AGCCATCATGAGTGAAGCTATTCAGAAAGCAATCAGAGAAATTGGAAGCTGATGAG 158
QY      32  AspLeuIleIleSerIleLeuAlaHisArgSerAlaGluGlnArgValIleArgGln 51
      159  AAATGCTCATCGACATTCGACTGAGAGAGCTCAATGCACAGCGGACGCTGATTGTAAG 218
QY      52  AlaTyHisGluThrTyGlyGluAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 71
      219  GAATATCAAGCAGCAGATATGAAAGAGCTGAAAGATGACTTGAAGGAGTCTCTCGGC 278
QY      72  AspPheGluArgAlaIleLeuLeuTrpThrLeuGluProGlyGlnArgAlaLeuLeu 91
      279  CACTTGAAGCATCTCATGAGGCGCTGAGCTCCACCGACGACTTTGATGCAAAAGCAG 338
QY      92  AlaAsnGluAlaThrIlyArgTrpThrSerSerAsnGlnValLeuMetGluValAlaCys 111
      339  CTAAGAAATCCATGAAGGCGGGGGAACAAGAGAGCCCTTGATGAAATCTTAATCT 398
QY      112  ThrArgThrSerThrGlnLeuLeuHisAlaArgGlnAlaTyHisAlaArgTyIlyLys 131
      399  ACCAGAGCAAGCAGGAAATGAGAGATATCTCTCAAGCCTATATACAGATACAGAGAG 458
QY      132  SerLeuGluGlnArgValAlaHisThrThrGlyAspPheArgIlyLeuLeuValSer 151
      459  AGCTTGAGATATCATAGTCCGAAACATCTGGTGACTCTCCGAAAGCTGTTGACT 518
QY      152  LeuValThrSerTrpArgTrpGlyGluGlyAspGluValAsnMetThrLeuAlaLysGln 171
      519  TTGGCAGATGGCAGAGAGATGAAGTCTGAAGTGAAGATGTCGCCAAACAAGAT 578
QY      172  AlaLysLeuValHisGluLysIleLysAspLysHisTyTrpAsnArgLysLys--ValIle 190
      579  GCCCAGATTCTCTATTAAGCTGGTGAACAGATGGGGCAGCGATGAAACAATTCAGCT 638
QY      191  ArgIleLeuSerThrArgSerLysAlaGlnIleAsnAlaThrPheAsnArgTyGlnAsp 210
      639  GAGATCTCTGTTTAAGAGCTTCTCTCAATTAAACATAACATTGATGAATACAGAAAT 698
QY      211  AspHisGlyGluGlnIleLeuLysSerLeuGluGluGlyAspAspAspLysPheLeu 230
      699  ATCAGCCAAAGACATTTGTCAGACATTA--AAAGAGAAATTAATCTGGGCAATTTTGA 755
QY      231  AlaLeuLysArgSerThrIleGlnLysLeuThrArgProGlnLeuTrpPheValAspVal 250
      756  GACTTACTGTGGCCATAGTATTAATGTGTGAGAAACACGCCGCCCTTTTAAGCCGAAAG 815
QY      251  LeuArgSerAlaIleAsnLysThrGlyThrAspGluGlyAlaLeuThrArgIleValTrp 270
      816  CTGCATCGAGCCCTTGAAGGCTATTCGAACGATGAGATTACTTCAACCGAATTAAGGTG 875
QY      271  ThrArgAlaGluIleAspLeuLysValIleGlyGluGlyTrpGlnArgArgAsnSerIle 290
      876  TCCAGATCAGAAATTAATCTTTGGACATTCGACAGAGATTCAGAGCAATTAATGCTCAT 935
QY      291  ProLeuGluLysAlaIleThrIlyAspTrpArgGlyAspTyGlyLysMetLeuValAla 310

```

```

DB      936  TCCCTATATTCAGCAATTAAATCGGATATCTTCGAGACTATGAATACACTTTTAAA 995
QY      311  LeuLeuGlyGluAspAsp 316
DB      996  ATCTGTGTGAGAGATGAC 1013

RESULT 6
US-09-023-655-1149
; Sequence 1149, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cooke, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; APPLICATION NUMBER:
; PRIOR APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1149:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1339 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9186967
; US-09-023-655-1149

Alignment Scores:
Pred. No.:      8.47e-50      Length:      1339
Score:          472.00      Matches:      111
Percent Similarity: 56.86%      Conservative: 63
Best Local Similarity: 36.27%      Mismatches: 130
Query Match:      29.30%      Indels:      2
DB:               4          Gaps:      2

US-10-690-564-2 (1-317) x US-09-023-655-1149 (1-1339)
QY      12  AlAProSerAspAlaGluGlnLeuArgThrAlaPheGluGlyTrpGlyThrAsnGln 31
      101  AGCCATCATGAGTGAAGCTATTCAGAAAGCAATCAGAGAAATTGGAAGCTGATGAG 160
QY      32  AspLeuIleIleSerIleLeuAlaHisArgSerAlaGluGlnArgValIleArgGln 51
      161  AAATGCTCATCGACATTCGACTGAGAGAGTCAATGACACAGCGGACGCTGATTGTTAAG 220

```

Oy 52 AATATGTAAGCAGCAATATGTAAGCAAGAGCTGAAAGATGACTGTAAGGCTGATCTCTGGC 280  
 Db 221 GAATATCAAGCAGCAATATGTAAGCAAGAGCTGAAAGATGACTGTAAGGCTGATCTCTGGC 280  
 Oy 72 ACPPhEgIuAArgAlaIleLeuLeuTprThrLeuGluProGlyGluuArgAspAlaLeuLeu 91  
 Db 281 CACTTTGAGCATCTCATGGTGGCCCTTAAGTACCTCCACCGACAGCTTTGTATGCAAAAGCAG 340  
 Oy 92 AlaAeNgIuAlaThrLysArgTrpThrSerSerAsnGluValLeuMetGluValaLys 111  
 Db 341 CTAAAGAAATTCATGAAAGCGCGGGAGCAAAAGAAAGATGCCCTTGATTTGAATCTTAAC 400  
 Oy 112 ThrArgThrSerThrGluLeuLeuHisValaArgGluAlaTyHisAlaArgTyLysVal 131  
 Db 401 ACCAGAGACAAAGCAGGCAAGATGAAGGATATCTTCGAAGCCATATATACAGTATCAAGAAAG 460  
 Oy 132 SerLeuGluGluAspValAlaHisIleThrThGluAspPheArgLysLeuLeuValSer 151  
 Db 461 AGCTCTTGAGATGACATTAAGTCCGAACAACATCTGGAGACTCCGGAAGCTCTGTGACT 520  
 Oy 152 LeuValThrSerTyArgTyArgGluAspGluValaAsnMetThrLeuAlaLysGluVal 171  
 Db 521 TTGGCAGAGTGGCAGAAAGATGAATACTCGAAAGTGATGAGAGATCTGGCCAAACAAGAT 580  
 Oy 172 AlaLysLeuValHisGluLysIleLysAspLysHisTyArgAsnArgLysPhe--ValIle 190  
 Db 581 GCCCAGATTCCTCATTAATGCTGAGTGAACAGATGGGCAAGATGAGACAAATTCACACT 640  
 Oy 191 ArgIleLeuSerThrArgSerLysAlaGlnIleAsnAlaThrPheAsnArgTyArgLys 210  
 Db 641 GAGATCTCTGTGTTAAGAGCTTTCTCCAAATTAACCTAATGATGATACAGAAAT 700  
 Oy 211 AspHisGlyGluIleLeuLysSerLeuGluGluValAspAspAspAspLysPheLeu 230  
 Db 701 ATCAGACCAAAAGACATTTGTGACAGCAGCA--AAAGAGAAATATCTGGCCATTTTGAA 757  
 Oy 231 AlaLeuLeuArgSerThrIleGlnCysLeuThrArgProGluLeuTyPheValAspVal 250  
 Db 758 GACTTATCTGTGGCCATGATTAAATGTGTGAGGAACAGCCGGCTTTTAAAGCCGAAAG 817  
 Oy 251 LeuArgSerAlaIleAsnLysThrGlyThrAspGluValAlaLeuThrArgIleValThr 270  
 Db 818 CTGCATCGAGCTTGAAAGGTATTTGGAACGTATGATTTACTCTTGAAACGAAATAATGCTG 877  
 Oy 271 ThrArgAlaGlnIleAspLeuLysValIleGlyGluGluTyGlnArgArgAsnSerIle 290  
 Db 878 TCCAAGTCAAGAAATTGACCTTTTGGACATTTCCGAACAGAGTTCAAGAAAGCATTAAGCTCAT 937  
 Oy 291 ProLeuGluLysAlaIleThrLysAspThrArgGlyAspTyGluLysMetLeuValAla 310  
 Db 938 TCCCTATATTCAGCAATTAAATCGATTACTTTCGAGAGCATGAATCAACACTTTAAA 997  
 Oy 311 LeuLeuGlyLysAspAsp 316  
 Db 998 ATCTGTGTGGAGATGAC 1015  
 RESULT 7  
 US-09-949-016-722  
 ; Sequence 722, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949, 016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08

	NUMBER OF SEQ. ID NOS: 207012	SOFTWARE: FastSeq for Windows Version 4.0
;	SEQ ID NO 722	;
;	LENGTH: 1339	;
;	TYPE: DNA	;
;	ORGANISM: Human	;
US-09-949-016-722		
Alignment Scores:		
Pred. NO.:	8,47e-50	Length: 1339
Score:	472.00	Matches: 111
Percent Similarity:	56.86%	Conservative: 63
Best Local Similarity:	36.27%	Mismatch: 130
Query Match:	29.30%	Indels: 2
DB:	4	Gaps: 2
US-10-690-564-2 (1-317) x US-09-949-016-722 (1-1339)		
QY	12	AlAProSeRAspApLaGlInLeuYrThrAlAphEgLuGlYTrPGLYThAsnGLu 31
DB	101	AGCCCACTCAGTGATGCTGAACTTATTCAGAAAGCAATCAGAGAAATTGCAATGATGAG 160
QY	32	AsPLeuLeIeSeRtLeuLaIhIsrGSeRAlaGlInaRgYbValIleArgIn 51
DB	161	AAATGCTCATCGCACTTCTGACTGAGAGGTCMAATGCAACGCGCGCATGATTTGTTAA 220
QY	52	AlAeYrHiEgLuThrYrGlYGLuApeLeuLeuYsThLeuApeRlyGLuLeuSeRAsn 71
DB	221	GAATATTCAGACACATATATGAAAGAGCTGAAAGATGACTTGAAGGGTATCTCTGGC 280
QY	72	AsPheGLuARgAlaIleLeuLeuTrPThLeuGLuProGLYGLuARgAsPaIeLeuLeu 91
DB	281	CACTTTGAGCATCTCATGTGTGGCCCTAGACATCCACGACAGATCTTTGATGCAAAAGCAG 340
QY	92	AlAAsnGLuAlAthrLYeArGrTrPThSeRSeRsnGLInValLeuMeGLuAlAlAcys 111
DB	341	CTAAAGAAATCCATGAAAGGCGCGGAAACAAACGAAGATGCCCTGATTTGAATTTAACT 400
QY	112	ThrArgThSeRtGLInLeuLeuHIsAlaArgInAlAYrHIsAlaArgTYrLYs 131
DB	401	ACCAAGCAACAGCGCAATGAAAGATATCTCTCAAGCCCTTTATACAGTATACAGAG 460
QY	132	SeRLeuGLuLaRPaValaHIsHIsrThrThrGLyAsPheARgLYsLeuLeuValSer 151
DB	461	AGCTTTGAGATGACATTTAGTCCGAAACATCTGGTGACTTCGGAAGACCTCTGTACT 520
QY	152	LeuValThSeRtYrGrTYrGLuGLyApeGLuValAeMeThrLeuAlaYbGLIn 171
DB	521	TTGGCAGATGCGCAAGAGATGAAAGTGTGAAAGTGATGCGATCTCGGCCCAACAAAGAT 580
QY	172	AlAlyLeuValHIsGLuYsYrLeuApeRlySHIsrYrAaNaRgLuApe--ValIle 190
DB	581	GCCCAAGATTTCTATTAAGCTGTGTAAGAACAGATGGGACGAGATGAAGAAATTCACT 640
QY	191	ArgIleLeuSeRtHrArgSeRtYsAlaGLInIleAsnAlaThrPheAsnARgTYrGLInAsp 210
DB	641	GAAATCTGTGTTTAAAGACTTCTCTCATTTAAACATAATTTAGATGATACAGAAAT 700
QY	211	AsPHeGLyGLuGLuIleuLYsSeRLeuGLuGLyAsPaRPaAsPaRPaRlyPheLeu 230
DB	701	ATCAGCCAAAGGACATGTGACAGCATTA--AAAGGAAATATCTGGGCATTTTGAA 750
QY	231	AlALeuLeuARgSeRtHrIleGLInYsLeuThrARgProGLuLeuYrPheValAsPaVal 250
DB	758	GACTTACTGTGGCCATTAAGTTAATGTGTGAGAAACACGCGGCTTTTAAAGCCGAAGA 817
QY	251	LeuARgSeRAlAlEeSnYrThrGLYHrApeGLuGLyAlaLeuThrArgIleValThr 270
DB	818	CTGCATCGACCTTGAAGGTATTGGAACTGATGATTTACTCTGAACCGAATATATGCTG 877
QY	271	ThrArgAlaGLuIleApeLeuYsValIleGLyGLuYrGLInARgARgAsnSerIle 290
DB	878	TTCAAGATTCGAATTTGACCTTTTGACATTTGGAACAGGTTTCAAGAAAGCTTATGCTAT 937



[illegible][illegible]

```

Db      1470 GAGAGGCTCAACAAGGCCATGATGGGGGCGCAGAGAAACAAGACCCGACCTGATTGGCATC 1529
Qy      269 ValTtRtRrAgaAlaGluIleAspLeuYsValIlegIyGluGluTyTgLnArGaRghsn 288
      :+:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::
Db      1530 ATGGTCTCTCCGACGAGACCGACCTCTCGACACATGCAGATCATATACCGGATGTAC 1589
Qy      289 SerIleProLeuGluTyLysAlaIleThrIlyAspThrArGgiYaaPyTyTgLyMetLeu 308
      |||:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::
Db      1590 GCGAAGTCGCTGTACCAACGACATCTCCGGGAAGTACTTCAAGGGGATTCACGGAAGATTCTG 1649
Qy      309 ValAlaLeuLeuGluIyGluAspAsp 316
      :::::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::
Db      1650 CTGAAGATCTGTGTGGCGCAATGAC 1673

RESULT 11
US-08-526-136-1
/ Sequence 1, Application US/08526136
/ Patent No. 6107089
/ GENERAL INFORMATION:
/ APPLICANT: Towle, Christine A. et al.
/ TITLE OF INVENTION: ANNEXIN XI
/ NUMBER OF SEQUENCES: 36
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: U.S.A.
/ ZIP: 02110-2804
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ COMPUTER: IBM PS/2 Model 502 or 55SX
/ OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
/ SOFTWARE: WordPerfect (Version 5.0)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/526.136
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/214.036
/ FILING DATE:
/ APPLICATION NUMBER: 07/837.775
/ FILING DATE: February 13, 1992
/ APPLICATION NUMBER: 07/764,465
/ FILING DATE: September 23, 1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Clark, Paul T.
/ REGISTRATION NUMBER: 30,162
/ REFERENCE/DOCKET NUMBER: 00786/099001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 542-5070
/ TELEFAX: (617) 542-8906
/ TELEX: 200154
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2305
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-526-136-1

Alignment Scores:
Pred. No.: 1.25e-46 Length: 2305
Score: 450.00 Matches: 111
Percent Similarity: 54.55% Conservative: 57
Best Local Similarity: 36.04% Mismatches: 132
Query Match: 27.93% Indels: 8
DB: 3 Gaps: 3

US-10-690-564-2 (1-317) x US-08-526-136-1 (1-2305)
Qy      13 ProSerAspAspAlaGluGluLeuArGThrAlaPheGluGluTyTgPyGlyThrAsnGluAsp 32
      |||||:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::
      :+:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::++:::

```

Db 699 CCCGTCGAGATGCTGAAGTCTCGTGGGAAAGGCATGAAGGGCTTTGGGACTGACGAGCG 758  
 QY 33 LeuIleIleSerIleLeuAlaHisArgSerAlaGluGlnArgLysValIleArgGlnAla 52  
 Db 759 GCCATCATTTGACCTGCCCTGGTATGTCGCTCCAAACAAGCAACGACAGATCTCTCTGCG 818  
 QY 53 TyrHisGluThrTyrGlyGluAspLeuLeuLysThrLeuAspLysGluLeuSerAsnAsp 72  
 Db 819 TTCAAGACACGATATGGAGAGGATTGTGATCAAAAGATCTGAAATCTGAACTGTCAAGAAAC 878  
 QY 73 PheGluArgAlaIleLeuLeuLeuTyrThrLeuGluProGlyGluArgAspAlaLeuLeuAla 92  
 Db 879 TTTGAGAGACATATCTTGCCCTGTATGAAGACCCCTGTCTCTTTGACGCTTATAGATA 938  
 QY 93 AsnGluAlaThrLysArgTyrThrSerSerAsnGlnValIleMetGluValAlaCysThr 112  
 Db 939 AAGGAAGATCAACAGGGGCGGCGCACTGATGAAGCTCCCTGATGAAGATCTCGGCTCC 998  
 QY 113 ArgThrSerThrGlnLeuLeuLeuHisAlaIleArgGlnAlaTyrHisAlaArgTyrLysLysSer 132  
 Db 999 CGCAGCAACGAGCACATCCGGGAGCTGAACAGAGTCTACAGACAGAAATCCAAAAAGACC 1055  
 QY 133 LeuGluGluAspValAlaHisIleThrThrGlyAspPheArgLysLeuLeuValSerLeu 152  
 Db 1059 CTGAGAGAGGCCATTCGAGGAGCACCTTCAAGGCACCTTCAAGCGCTCTCATCTCTCTC 1111  
 QY 153 ValThrSerTyrArgTyrGlyGluLysAspLysValAsnMetThrLeuAlaLysGlnLysAla 172  
 Db 1119 TCTCAGGAAACCGGGATGAAAGCAACAACGTGACATGACCTTGTCCAGAGAGATGTG 1176  
 QY 173 LysLeuValHisGluLysIleLysAspLysHisIleTyrAsnAspGluAspValIleArg--- 191  
 Db 1179 CAGAGAGCTCTATGACGCTGGGAGAGAACCGCTGGGACAGATGATCCAGTTCAGTCCG 1236  
 QY 192 IleLeuSerThrArgSerLysAlaGlnIleAsnAlaThrPheAsnArgTyrGlnAspAsp 211  
 Db 1239 ATTCTGTGCTCCCGGAGCGCGGCCACCACTGGTGCGCATTTTAAAGATATCAAGAGATG 1299  
 QY 212 HisGlyGluGlnIleLeuLysSerLeu-----GluGluGlyAspAspAspAspLys 228  
 Db 1299 ACAGGACGTGACATTTGAGAAGAGCATCTCGCGGAGATGTCCGGGACCTGGAGAGAGGCC 1355  
 QY 229 PheLeuAlaLeuLeuArgSerThrIleGlnCysLeuThrArgProGluLeuTyrThrVal 248  
 Db 1359 ATGCTGGCTGTG-----GTGAAATGTCCTTAAAGATACCCAGCCCTCTTTGCT 1400  
 QY 249 AspValLeuArgSerAlaIleAsnLysThrGlyThrAspGluGlyAlaLeuThrArgIle 268  
 Db 1407 GAAGAGCTTCACAAAGGCCATGGGGAGCCGGACCAAAAGACCGGACCTGTATCCGCAATC 1466  
 QY 269 ValThrThrArgAlaGluIleAspLeuLysValIleGlyGluGluTyrGlnArgArgAsn 288  
 Db 1467 ATGGTGTCTCCAGAGGATGACCTCTCGACATCAAGCAGAGATTAAGCGGTGAT 1522  
 QY 289 SerIleProLeuGluLysAlaIleThrLysAspThrArgLysPyrGlyLysMetLeu 308  
 Db 1527 GCGAAGTCGTGTACCAAGCATCAACGGGAGACACTTCAGGGGATTAACGGAAAGATCTG 1586  
 QY 309 ValAlaLeuLeuGlyGluAspAsp 316  
 Db 1587 CTGAAGATCTGTGGCAATGAC 1610  
 RESULT 12  
 US-08-526-136-3  
 ; Sequence 3, Application US/08526136  
 ; Patent No. 6107089  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Towle, Christine A. et al.  
 ; TITLE OF INVENTION: ANNEXIN XI  
 ; NUMBER OF SEQUENCES: 36  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson  
 ; STREET: 225 Franklin Street



```

Db      506 GATGCAAAATTCCTGCTAGCAAGCAATGAGGCTTTTGGACAGATGACGAGCAATGTG 565
Qy      36 Ser11eLeuAlaHisIserSerAlaGluGlnArgVal11eArgGlnAlaTyRHisGlu 55
Db      566 GATGTGGTGGCCAAACCGTTCATATGATGACAGGCAAAAATAAGAGCATTTAAAGACC 625
Qy      56 ThTyGlyGluAspLeuLeuTyThrLeuAspLysGluLeuSerAsnAspPheGluArg 75
Db      626 TCCTATGGCAAGATTTATCAAAAGATCTCAATCAGATTAAGTGAATAATGGAAGA 665
Qy      76 Ala11eLeuLeuTyThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAla 95
Db      686 CTGATCTCGCCCTTCATGCTCCCTCACTATTACAGATCCGAGCTTAACGAAAGCA 745
Qy      96 ThTyLeuArgTyThrIserSerAsnGlnValLeuMetGluValAlaCysThrArgThrSer 115
Db      746 ATGCAAGGAGACAGAACTCAGAACTGTAATTAATGAGATTTTGTGCACAAAGCAAAAT 805
Qy      116 ThGlnLeuLeuHisAlaArgGlnAlaTyRHisAlaArgTyRHisLysSerLeuGluGlu 135
Db      806 CAGAAATCCGAAATTTGTCAGATGTTATCAGTCAAGAAATTTGAGCAGAGCCTTGAAAG 865
Qy      136 AspValAlaHisIserThrGlyAspPheArgLysLeuLeuValSerLeuValThrSer 155
Db      866 GACATTAGGTCAGATTCATCAGACATTTTGAACGTTTACTGTGTCCATGTCACAGGA 925
Qy      156 TyArgTyGluGluArgLysArgValAsnMetThrLeuAlaGlnGlnAlaLysLeuVal 175
Db      926 AATGCGATAGAAACCAAGATAAACCACCAATAGCTCAGAAAGTGTCCAGCCCTTC 985
Qy      176 HisGluLys11eLysAspLysHisTyRAsnAspGluAspVal11eArg--11eLeuSer 194
Db      986 TATCAAGCTGTGAGGGGAGACCTAGGAGCGATGAAATCTTGCTTAACATGACTCTTGC 1045
Qy      195 ThrArgSerLysAlaGln11eAsnAlaThrPheAsnArgTyRGlutAspHisGluGlu 214
Db      1046 ACAAGAGAGCTTCTCCTAGCTGAGAGCTACCATGAGGCTTATCTGAGATGCTAATCGA 1105
Qy      215 Glu11eLeuLysSer-----LeuGlnGluLysAspAsp 226
Db      1106 GACTGTGTAAGCGCTGAGCCCTGAGCTTCCGATATGTAGAAAGTGT----- 1156
Qy      227 AspLysPheLeuAlaLeuLeuArgSerThr11eGlnCysLeuThrArgProGluLeuTyR 246
Db      1157 -----TTGAAGACCATCTTCAGATGTCCCTGAACCGCCCTGCTTC 1198
Qy      247 PheValAspValLeuArgSerAla11eAsnLysThrGlyThrAspGluGluAlaLeuThr 266
Db      1199 TTGTGAGAGGCTCTACTATGCTATGAAAGGTGCTGGCAGATGATCCACCCCTGCTC 1258
Qy      267 Arg11eValThrThrArgAlaGln11eAspLeuVal11eGluGlnGluTyRGlutAsp 286
Db      1259 CGGATGTGTCTCCTGAGAGTGAATGATCTGTCAAAATAAACAGATGTGCTGCTCAG 1318
Qy      287 ArgAsnSer11eProLeuGlnLysAla11eThrLysAspThrArgLysLysArgTyRGlut 306
Db      1319 ATGTATCAAGAAAGCTGCGGCAACAATGATTCAGAGTGAACGAGTGAATTCAGAAAG 1378
Qy      307 MetLeuValAlaLeuLeuGluGlu 314
Db      1379 CTCTCTCTGCTGCTATGTGTGGCCAG 1402

RESULT 14
US-09-949-016-1200
; Sequence 1200, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

```

```

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1200
; LENGTH: 2054
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1200

Alignment Scores:
Score: 2.51e-46 Length: 2054
Percent Similarity: 447.00 Matches: 102
Percent Local Similarity: 54.874 Conservative: 67
Query Match: 33.124 Mismatches: 121
DB: 27.754 Indels: 18
Gaps: 3

US-10-690-564-2 (1-317) x US-09-949-016-1200 (1-2054)
Qy      16 AspAlaGluGlnLeuArgThrAlaPheGluGlyTyRArgLysAsnGluAspLeu1e1e 35
Db      506 GATGCAAAATTCCTGCTAGCAAGCAATGAGGCTTTTGGACAGATGACGAGCAATGTG 565
Qy      36 Ser11eLeuAlaHisIserSerAlaGluGlnArgVal11eArgGlnAlaTyRHisGlu 55
Db      566 GATGTGGTGGCCAAACCGTTCATATGATGACAGGCAAAAATAAGAGCATTTAAAGACC 625
Qy      56 ThTyGlyGluAspLeuLeuTyThrLeuAspLysGluLeuSerAsnAspPheGluArg 75
Db      626 TCCTATGGCAAGATTTATCAAAAGATCTCAATCAGATTAAGTGAATAATGGAAGA 685
Qy      76 Ala11eLeuLeuTyThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAla 95
Db      686 CTGATCTCGCCCTTCATGCTCCCTCACTATTACAGATCCGAGCTTAACGAAAGCA 745
Qy      96 ThTyLeuArgTyThrIserSerAsnGlnValLeuMetGluValAlaCysThrArgThrSer 115
Db      746 ATGCAAGGAGACAGAACTCAGAACTGTAATTAATGAGATTTTGTGCACAAAGCAAAAT 805
Qy      116 ThGlnLeuLeuHisAlaArgGlnAlaTyRHisAlaArgTyRHisLysSerLeuGluGlu 135
Db      806 CAGAAATCCGAAATTTGTCAGATGTTATCAGTCAAGAAATTTGAGCAGAGCCTTGAAAG 865
Qy      136 AspValAlaHisIserThrGlyAspPheArgLysLeuLeuValSerLeuValThrSer 155
Db      866 GACATTAGGTCAGATTCATCAGACATTTTGAACGTTTACTGTGTCCATGTCACAGGA 925
Qy      156 TyArgTyGluGluArgLysArgValAsnMetThrLeuAlaGlnGlnAlaLysLeuVal 175
Db      926 AATGCGATAGAAACCAAGATAAACCACCAATAGCTCAGAAAGTGTCCAGCCCTTC 985
Qy      176 HisGluLys11eLysAspLysHisTyRAsnAspGluAspVal11eArg--11eLeuSer 194
Db      986 TATCAAGCTGTGAGGGGAGACTAGGACCGAGTAATCTTGCTTTAATGATGCTTGTGC 1045
Qy      195 ThrArgSerLysAlaGln11eAsnAlaThrPheAsnArgTyRGlutAspHisGluGlu 214
Db      1046 ACAAGAGCTTCTCCTAGCTGAGAGCTACCATGAGGCTTATCTGAGATGCTAATCGA 1105
Qy      215 Glu11eLeuLysSer-----LeuGlnGluLysAspAsp 226
Db      1106 GACTGTGTAAGCGCTGAGCCCTGAGCTTCCGATATGTAGAAAGTGT----- 1156
Qy      227 AspLysPheLeuAlaLeuLeuArgSerThr11eGlnCysLeuThrArgProGluLeuTyR 246
Db      1157 -----TTGAAGACCATCTTCAGATGTCCCTGAACCGCCCTGCTTC 1198

```





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2005, 11:53:35 ; Search time 751 Seconds

(without alignments)  
2742.375 Million cell updates/sec

Title: US-10-690-564-2

Perfect score: 1611

Sequence: 1 MATLKXSDSVAPASDABDL.....KOTRGDYEKMLVALLSDDA 317

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10690564.rnuc 19082005.172309.25406/app.query.fasta\_1.455  
-DB=Published Applications NA -QFMT=fastcap -SUPERF=rimb -MINMATCH=0.1  
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=JOCL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10690564.@CGN 1.1 480 @rnuc 19082005.172309.25406  
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -MAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCR\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09C\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq:\*  
23: /cgn2\_6/ptodata/2/pubpna/US10J\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_PUBCOMB.seq:\*  
25: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1611	100.0	1230	21 US-10-690-564-1	Sequence 1, Appl
2	1202.5	74.6	1184	19 US-10-767-795-4242	Sequence 4242, Ap
3	1185.5	73.6	2588	17 US-10-393-840-28	Sequence 28, Appl
4	1177.5	73.1	1145	18 US-10-425-114-7439	Sequence 7439, Ap
5	1177.5	73.1	1145	18 US-10-425-114-29928	Sequence 29928, Ap
6	1177.5	73.1	1133	18 US-10-424-599-1190	Sequence 1190, Ap
7	1171.5	72.7	1293	14 US-10-219-220-17	Sequence 17, Appl
8	1171.5	72.7	1293	14 US-10-393-840-109	Sequence 109, Appl
9	1132.5	70.3	1058	14 US-10-219-220-218	Sequence 218, Appl
10	1112.5	69.1	1145	18 US-10-425-114-10772	Sequence 10772, A
11	1046.5	65.0	1231	18 US-10-424-599-70457	Sequence 70457, A
12	1045	64.9	635	9 US-09-770-149-655	Sequence 655, Appl
13	1011.5	62.8	1385	19 US-10-767-701-13525	Sequence 13525, A
14	1009.5	62.7	1235	18 US-10-425-114-21986	Sequence 21986, A
15	1009.5	62.7	1285	18 US-10-425-114-26803	Sequence 26803, A
16	1009.5	62.7	1972	20 US-10-425-115-2178	Sequence 2178, Ap
17	1008.5	62.6	1235	18 US-10-425-114-35285	Sequence 35285, A
18	1008.5	62.6	1243	18 US-10-425-114-23365	Sequence 23365, A
19	1004.5	60.4	1292	19 US-10-437-963-78307	Sequence 78307, A
20	975.5	60.6	1337	19 US-10-767-701-14545	Sequence 14545, A
21	975.5	60.6	1476	19 US-10-437-963-58743	Sequence 58743, A
22	951.5	59.1	1496	20 US-10-425-115-63693	Sequence 63693, A
23	949.5	58.9	1300	20 US-10-739-930-2276	Sequence 2276, Ap
24	919.5	57.1	908	18 US-10-425-114-10228	Sequence 10228, A
25	866.5	53.8	789	14 US-10-219-220-20	Sequence 20, Appl
26	866.5	53.8	789	17 US-10-393-840-26	Sequence 26, Appl
27	866.5	53.8	789	17 US-10-393-840-112	Sequence 112, Appl
28	832	51.6	1437	18 US-10-424-599-88585	Sequence 88585, A
29	825	51.2	1499	18 US-10-424-599-30062	Sequence 30062, A
30	793	49.2	589	19 US-10-021-323-13465	Sequence 13465, A
31	786.5	48.8	1153	18 US-10-425-114-7980	Sequence 7980, Ap
32	786.5	48.8	1172	18 US-10-425-114-13073	Sequence 13073, A
33	786.5	48.8	2336	18 US-10-424-599-137007	Sequence 137007, A
34	785	48.7	602	19 US-10-021-323-12115	Sequence 12115, A
35	784	48.7	1261	18 US-10-424-599-87476	Sequence 87476, A
36	782	48.5	1194	20 US-10-739-930-3735	Sequence 3735, Ap
37	775	48.1	591	19 US-10-021-323-15046	Sequence 15046, A
38	773.5	48.0	1320	18 US-10-424-599-137005	Sequence 137005, A
39	760	47.2	600	19 US-10-021-323-15621	Sequence 15621, A
40	759	47.1	1113	18 US-10-424-599-130903	Sequence 130903, A
41	757.5	47.0	2132	17 US-10-393-840-27	Sequence 27, Appl
42	742	46.1	618	19 US-10-767-795-4245	Sequence 4245, Ap
43	738	45.8	586	19 US-10-021-323-15526	Sequence 15526, A
44	728	45.2	2612	18 US-10-424-599-68473	Sequence 68473, A
45	691.5	42.9	870	19 US-10-437-963-18738	Sequence 18738, A

#### ALIGNMENTS

RESULT 1  
US-10-690-564-1  
; Sequence 1, Application US/10690564  
; Publication No. US20050089872A1  
; GENERAL INFORMATION:  
; APPLICANT: KOREA KUNHO PETROCHEMICAL CO., LTD  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ANNEXINS FROM PLANTS  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/690,564  
; CURRENT FILING DATE: 2003-10-23  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Seqpatentin 1.71  
; SEQ ID NO 1  
; LENGTH: 1230  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana

US-10-690-564-1

## Alignment Scores:

Pred. No.:	1,766-165	Length:	1230
Score:	1611.00	Matches:	317
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-10-690-564-2 (1-317) x US-10-690-564-1 (1-1230)

```
QY      1 MetAlaThrLeuIuYsValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20
DB      34 ATGGCGACTCTTAAGGTTTCTGAATTCCTGCTCGCTCTTCTGATGATGCTGAGCAATG 93
QY      21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
DB      94 AGAACCGCTTTTAAAGATGGGGTACGACAGAGACTTGTATCATCAATCAATCTTGCTCAC 153
QY      41 ArgSerAlaGluGluAspValIleArgAlaAlaTyrHisGluThrTyrGlyGluAsp 60
DB      154 AGAAGTCTTAACAGAGAAATGCTATCGACCAAGCATCAACAAACCTTACGGCGAAGAC 213
QY      61 LeuLeuYsThrLeuAspIuSerIuSerAspPheGluAlaIleLeuLeuTyr 80
DB      214 CTTCTCAAGACTCTTGAACAGAGACTCTTACGATTTGAGAGAGCTATCTTGTTGG 273
QY      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIlySAtgTTPTr 100
DB      274 ACTCTTGAAACCGGTGCGGTGATGCTTATTTGGCTAATACCTCAAAAAGATGACT 333
QY      101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGluLeuHis 120
DB      334 TCACGACACCAAGTCTTATGAGAACTGCTTGCACAGAGCAATCAACGCGCTCTTAC 393
QY      121 AlaArgGluAlaTyrHisAlaArgTyrIlySylSerLeuGluIuAspValAlaHis 140
DB      394 GTAGGCAAGCTTACCATGCTCGCTACAGAGATCTTGTAGAGAGAGCGTTGCTCACAC 453
QY      141 ThrThrGlyAspPheArgIlySylLeuLeuValSerLeuValThrSerTyrArgTyrGly 160
DB      454 ACTACCGGTGACTCGAAAGCTTTGGTTCTCTGTATCCCTCATACAGGTACAGAGGA 513
QY      161 AspGluValAsnMetThrLeuAlaIlySylGluAlaIleAspLeuValHisGluIlySyl 180
DB      514 GATGAAGTGAACATGACATTGGCTTACAGCAAGAACTTACGCTTCATGAGAAATCAG 573
QY      181 AspIlyHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerIlySylGln 200
DB      574 GACAAAGCACTACATGATGAGATGTTATTAAGATCTTGTCCACAAAGAACAAAGCTCAG 633
QY      201 IleAsnAlaThrPheAsnArgTyrGluAspAspHisGlyGluGluIleLeuIlySerLeu 220
DB      634 ATCAATGCTATCTTTAAACCTTACCAAGATGATCTGCGAGGAAATTTCCCAAGACTTT 693
QY      221 GluGluGlyAspAspAspAspIlySylPheLeuAlaLeuLeuArgSerThrIleGlnCyl 240
DB      694 GAGGAAGAGATGATGATGACAGTTCTTGCACATTTTGAGGTCAACCATTCAGTGGCTTG 753
QY      241 ThrArgProGluLeuTyrPheValAspValIleuArgSerAlaIleAsnIlySylGly 260
DB      754 ACAAGACCAAGAGCTTACTTGTGATGTTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 813
QY      261 AspGluGlyAlaLeuThrArgIleValIleThrThrArgAlaGluIleAspLeuIlySyl 280
DB      814 GATGAAGAGAGCTTACATTAATTTGACCAAGAGCTGAGATTGACTTGAAGAGTCAAT 873
QY      281 GluGluGlyTyrGluArgAspAsnSerIleProLeuGluIlySylAlaIleThrIlySyl 300
DB      874 GGAAGAGAGTACCAAGCGCAAGAACAGCATTCCTTGGAGAAAGCTTATTCACAAAGACT 933
QY      301 ArgGlyAspTyrGluIlySylMetLeuValAlaLeuLeuGlyIuAspAspAla 317
```

DB 934 CGTGAGATTACAGAGATGCTCGTGCACCTTCTGCTGAGAGATGATGCT 984

## RESULT 2

```
US-10-767-795-4242
; Sequence 4242, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yinhua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 4242
; LENGTH: 1184
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHT-09MAY01-C101_1
US-10-767-795-4242
```

## Alignment Scores:

Pred. No.:	6,1e-121	Length:	1184
Score:	1202.50	Matches:	230
Percent Similarity:	85.30%	Conservative:	37
Best Local Similarity:	73.48%	Mismatches:	45
Query Match:	74.64%	Indels:	1
DB:	19	Gaps:	1

US-10-690-564-2 (1-317) x US-10-767-795-4242 (1-1184)

```
QY      1 MetAlaThrLeuIuYsValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20
DB      70 ATGGCGACTCTTAAGGTTTCTGAATTCCTGCTCGCTCTTCTGATGATGCTGAGCAACTA 129
QY      21 ArgThrAlaPheGluGlyTTPGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
DB      130 AGAAGACCTTTTACAGATGGGAACTAATGAGGCTTATATCATATATATTTGGCTCAC 189
QY      41 ArgSerAlaGluGluAspValIleArgGluAlaTyrHisGluThrTyrGlyGluAsp 60
DB      190 AGAATGCCGAGCAAGAACTGATGAAAACTTACGCTGAACCTATGAGAGAT 249
QY      61 LeuLeuYsThrLeuAspIuSerIuSerAspPheGluAlaIleLeuLeuTyr 80
DB      250 CTCCTCAAGCACTGACAGAGAGCTCTCGAATGACTTGTGAGAGGCTGTTTGGCTTTGG 309
QY      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIlySylGlyTTPTr 100
DB      310 GCTCTTGATCTCGTGAAGCTGATGCTTTTGGCTTATGAGAGCCCAAAAGTGGACT 369
QY      101 SerSerAsnGluValLeuMetGluValAlaCysThrArgThrSerThrGluLeuHis 120
DB      370 TCACGACATCAGGTCCTTATGGAATAGCTTGCACAAAGTCTGCCAACCAACTGCTTGCAC 429
QY      121 AlaArgGluAlaTyrHisAlaArgTyrIlySylSerLeuGluIuAspValAlaHis 140
DB      430 GCAGGCAAGCTTATGATGCTGTTAAGAGTGCCTTGAAGAGATTTGCTCATCAC 489
QY      141 ThrThrGlyAspPheArgIlySylLeuLeuValThrSerTyrArgTyrGlyGly 160
DB      490 ACGACTGGGAGCTTCGTAGAGCTCTCTCACTTCACTAGTAGTCAACAGATATGAGGA 549
QY      161 AspGluValAsnMetThrLeuAlaIlySylGluAlaIleAspLeuValHisGluIlySyl 180
DB      550 GAGAGGTGAACATGAAATCTGGCGAAACAGAGCGCAATTTCTTATAGAAAAATTTCA 609
QY      181 AspIlyHisTyrAsnAspGluAspValIleArgIleLeuSerThrArgSerIlySylGln 200
```

Db 610 GACAAAGCTTACAGTATGACGATGTATAGAGGTTTGGCTACAGAGAGGACGAG 669  
Qy 201 ILeasnaIaThrPheAsnAspGlyTrGlnAspAspHisGlyGlnGluIleLeuLeuSerLeu 220  
Db 670 ATCAATGCACTCTGAAATACACTACAAATATGAAATATGAAATATGAAATATGAAATATG 729  
Qy 221 GlnGluGluAspAspAspAspPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
Db 730 ---AAGGTGATCTCTAAGAGATAGTCTCTGCACTACTAAGGTCTCAAGAGAGGCTGG 786  
Qy 241 ThrArgProGluLeuLeuLeuPheValAspValLeuArgSerAlaIleAsnLeuThrGlyThr 260  
Db 787 GTCTATCCGAAAGATATTTTGGAGAGGTTCTTCCGCTGCAATCAATAGAGAGAGAG 846  
Qy 261 AspGluGlnAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280  
Db 847 GATAGAGAGCTCTTACTAGAGTGTGTGCTAGGCTGAGGTTGATCTTAAAGATCTA 906  
Qy 281 GlnGluGluTrpGlnArgArgAsnSerIleProLeuGluValAlaIleThrIleAspThr 300  
Db 907 GCAGATGATACCAAGCAG 966  
Qy 301 ArgGlyAspTrpGlyLeuLeuMetLeuValAlaLeuLeuGly 313  
Db 967 CATGAGAGCTATGAAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005

RESULT 3  
US-10-393-840-28  
; Sequence 28, Application US/10393840  
; Publication No. US20030229922A1  
; GENERAL INFORMATION:  
; APPLICANT: Bloksberg, Leonard N.  
; TITLE OF INVENTION: Materials and Methods for the  
; FILE REFERENCE: 11000.1012C3  
; CURRENT APPLICATION NUMBER: US/10/393.840  
; PRIORITY FILING DATE: 2003-03-20  
; PRIORITY APPLICATION NUMBER: US 09/636,800  
; PRIORITY FILING DATE: 2000-08-10  
; PRIORITY APPLICATION NUMBER: US 09/170,862  
; PRIORITY FILING DATE: 1998-10-13  
; PRIORITY APPLICATION NUMBER: US 60/148,426  
; PRIORITY FILING DATE: 1999-08-11  
; PRIORITY APPLICATION NUMBER: PCT NZ/99/00169  
; NUMBER OF SEQ ID NOS: 956  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 2588  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-10-393-840-28

Alignment Scores:  
Pred. No.: 1,36-118 Length: 2588  
Score: 1185.50 Matches: 228  
Percent Similarity: 83.91% Conservative: 38  
Best Local Similarity: 71.92% Mismatches: 50  
Query Match: 73.59% Indels: 1  
DB: 17 Gaps: 1

US-10-690-564-2 (1-317) x US-10-393-840-28 (1-2588)

Qy 1 MetaIaThrLeuLeuValSerAspSerValProIaProSerAspAspAlaGluGlnLeu 20  
Db 1390 ATGTCGACTCTCAACGCTCCGACAGCACTGCCCCCTGAGCGATGATGAGAGCACTC 1449  
Qy 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40  
Db 1450 CGGACAGCTCTCGAGATGGGAAACAATGAGAGCTGATCAATCATATGCTGCTCAT 1509  
Qy 41 ArgSerAlaGluGlnArgLeuValIleArgGlnAlaTrpHisGluTrpTrpGlyGluAsp 60

Db 1510 AGGAATGGGGGCGAGAGAGAGCTGATTCGGCAAAACCTATGCGGAGCTTACGGGAGAG 1569  
Qy 61 LeuLeuLeuThrLeuAspLeuGluLeuSerAsnAspPheGlnAlaIleLeuLeuTrp 80  
Db 1570 CTCCTCAAGGCAATGGACAGAGAACTTACCAATGATTTCCAGAGGCTGGTGTCTTTGG 1629  
Qy 81 ThrLeuGluProGlyGlnArgAspAlaLeuLeuAlaAsnGluAlaThrIleValGlyTrpThr 100  
Db 1630 TCACTGATCGGCTGAAACCTGATGCGTACTTGGCCGATATACGAGAAAGATGGACT 1689  
Qy 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
Db 1690 TCAAGCAACAGGTTCTTAAGAAATAGCTTGCACAGAGCTCCGACAGAGTGTCTATG 1749  
Qy 121 AlaArgGlnAlaTrpHisAlaArgTrpLeuLeuSerLeuGluGluAspValAlaHisHis 140  
Db 1750 GCAAGACAAACATATCATGCCCCATCAAGAAAGTCCGAGAAAGAGAGAGAGAGAGAG 1809  
Qy 141 ThrThrGlyAspPheArgLeuLeuLeuValSerLeuValThrSerTrpArgTrpGluGly 160  
Db 1810 ACAACTGAGATTTTGTAAAGTTGCTGTGACTCTTGTGAGCTCTTACCGTTATGATGA 1869  
Qy 161 AspGluValAsnMetThrLeuAlaLeuGlnGluAlaLeuLeuValHisGluValIleLeu 180  
Db 1870 GATGAGCTGAATATGACTTTGGCAAAAGCAGAGGCTTAAGTACTCCAGAGAGATCTCA 1929  
Qy 181 AspLeuHisTrpAsnAspGluAspValIleArgIleLeuSerThrArgSerTrpAlaGln 200  
Db 1930 GAGAAAGCTTATGCGCATGAGATCTCATAGAGATTTTGGCTCTAGAGCAAAAGCAG 1989  
Qy 201 ILeasnaIaThrPheAsnAspGlyTrGlnAspAspHisGlyGlnGluIleLeuLeuSerLeu 220  
Db 1990 GTCAAATGCTTACGGTGAATCACTCAAAAAGAGTTGGAATATATATCAACAAGATTTG 2049  
Qy 221 GlnGluGluAspAspAspAspPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
Db 2050 AAAACT---GATCCAAAGAGCGGTTCTTACTACTAGAGACTACAGTAAAGTGGCTG 2106  
Qy 241 ThrArgProGluLeuLeuLeuPheValAspValLeuArgSerAlaIleAsnLeuThrGlyThr 260  
Db 2107 ACTGCGCCTGAGAAAGATATTTGAAAAGTTCTTGTGCTGACCATATAGAGAGAGACA 2166  
Qy 261 AspGluGlnAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuValIle 280  
Db 2167 GATGAAGGGGCTCTGACAGAGATGTTGCTTACAGAGGCGGAGGTTGACATGAAGTTTATA 2226  
Qy 281 GlnGluGluTrpGlnArgArgAsnSerIleProLeuGluValAlaIleThrIleAspThr 300  
Db 2227 AGTGAAGAGTACCAAG 2286  
Qy 301 ArgGlyAspTrpGlyLeuLeuMetLeuValAlaLeuLeuGlyGluAspAspAla 317  
Db 2287 ACTGAGAGCTATGAAAAATGCTTCTGCAATGATTTGGCAGCTGAGGCT 2337

RESULT 4  
US-10-425-114-7439  
; Sequence 7439, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; PRIORITY FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 7439  
; LENGTH: 1145

```

; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700661911_FLI
US-10-425-114-7439

```

## Alignment Scores:

```

Pred. No.:      3,11e-118      Length:      1145
Score:          1177.50      Matches:      226
Percent Similarity: 84.54%      Conservative: 42
Best Local Similarity: 71.29%      Mismatches: 48
Query Match:     73.09%      Indels:      1
DB:              18          Gaps:      1

```

US-10-690-564-2 (1-317) x US-10-425-114-7439 (1-1145)

```

QY      1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
DB      54 ATGGCGACACTTAAGGTTCTTCAGCCCTCTCCCGCTGGACAGATGATGTGACAGCTC 113
QY      21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
DB      114 CGAAAGACCTTTCAGAGTTGGGAACTAACAGAGAGCTGATGTATGATCTTGCTCAT 173
QY      41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrrHisGluThrTyrglyLysAsp 60
DB      174 AGGAATGCTGCTCAGAGGAAGCTTATCAGAAACTTATGCCAGACTTATGGAGAAGAT 233
QY      61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluLysAlaIleLeuLeuTrp 80
DB      234 CTCCTCAAGGCTTGGACAAAGAACTCAAGATGATTTTGAGAGGCTGCTCATCTTTGG 293
QY      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGlnAlaThrLysArgTrpThr 100
DB      294 ACCTTGATCTCTGCTGAAGCTGATGCTATTTTGGCAACAGCGACTAAAAATGAGCT 353
QY      101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
DB      354 TCAAGCATCAGGTTCTGATGAAATAGCTGCATAGCTTGAAGCTTGAACATTTGTTGCT 413
QY      121 AlaArgGlnAlaTyrrHisAlaArgTyrrLysLysSerLeuGluGluAspValAlaHisHis 140
DB      414 GCGAGAAAGCTTACCAATGTTCTTATTAAGAGCTCTCGAGAGAGATGTTGCTCATCAC 473
QY      141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrrArgTyrglyLys 160
DB      474 ACAACGAGAACTTCCGTAAGCTCATATCACTCTGCTTATGTTCTTATTCATATGAAGA 533
QY      161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
DB      534 GATGAGGTCAACTTGACCTGGCAAAAACGAGGCAAAATTTGCTGATGAGATTTTCA 593
QY      181 AspLysHisTyrrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
DB      594 AACAAAGCTTATATATGATGACACTTCATCAGATTTTGGCCACAAGGAGCAAGACAG 653
QY      201 ILeAsnAlaThrPheAsnArgTyrrGlnAspAspHisGlyGluGlnIleLeuLysSerLeu 220
DB      654 ATTAAATGCTACTTGAATCACTCAAAAGATGCAATTGGACAAAGATTCACAAAGAGACT 713
QY      221 GluGluGlyAspAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
DB      714 ---AAGGCTGATCCAAAGAGAGAGGTTCTGCTTTCCTTAAGAGCACTGGAAGTCTGG 770
QY      241 ThrArgProGluLeuTyrrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260
DB      771 ATCCGCTCCGAGAAAGAACTTCGAAAGGTGTTCGCTGCGATCAACAAGCGAGAGAAC 830
QY      261 AspGluGlyAlaLeuThrArgIleValIleThrArgAlaGluIleAspLeuLysValIle 280
DB      831 GATGAAGAGACCTTACAAAGATGAGTGGCCACAAGGCTGAGGTGATTTGAAAAACATA 890
QY      281 GlyGluGlyLysGlnArgAsnSerIleProLeuGluLysAlaIleThrLysAspThr 300

```

```

DB      891 GCAGATGAGTATCAAGAGAGAGAGAGTGCCTCTTGAACGTGATTTGTAAAGACACC 950
QY      301 ArgGlyAspTyrrGluLysMetLeuValAlaLeuLeuGlyLysAspAspAla 317
DB      951 ACCGTGACTATGAGAAATGCTTGTGGGCTTTTGAAGCATGATGATGCT 1001

```

## RESULT 5

US-10-425-114-29928

```

; Sequence 29928, Application US/10425114
; Publication No. US20040034808A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 29928
; LENGTH: 1145
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-GMROPIC033E01_FLI
US-10-425-114-29928

```

## Alignment Scores:

```

Pred. No.:      3,11e-118      Length:      1145
Score:          1177.50      Matches:      226
Percent Similarity: 84.54%      Conservative: 42
Best Local Similarity: 71.29%      Mismatches: 48
Query Match:     73.09%      Indels:      1
DB:              18          Gaps:      1

```

US-10-690-564-2 (1-317) x US-10-425-114-29928 (1-1145)

```

QY      1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
DB      54 ATGGCGACACTTAAGGTTCTTCAGCCCTCTCCCGCTGGACAGATGATGTGACAGCTC 113
QY      21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
DB      114 CGAAAGACCTTCTCAGTTGGGAACTAACAGAGAGCTGATGTATCGATCTTGGCTCAT 173
QY      41 ArgSerAlaGluGlnArgLysValIleArgGlnAlaTyrrHisGluThrTyrglyLysAsp 60
DB      174 AGGAATGCTGCTCAGAGGAAGCTTATCAGAAACTTATGCCAGACTTATGGAAGAT 233
QY      61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluLysAlaIleLeuLeuTrp 80
DB      234 CTCCTCAAGGCTTGGACAAAGAACTCAAGATGATTTTGAAGGCTGCTCATCTTGG 293
QY      81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGlnAlaThrLysArgTrpThr 100
DB      294 ACCTTGATCTCTGCTGAAGCTGATGCTATTTTGGCAACAGCGACTAAAAATGAGACT 353
QY      101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuLeuHis 120
DB      354 TCAAGCATCAGGTTCTGATGAAATAGCTGCATAGCTTGAAGCTTGAACATTTGTTGCT 413
QY      121 AlaArgGlnAlaTyrrHisAlaArgTyrrLysLysSerLeuGluGluAspValAlaHisHis 140
DB      414 GCGAGAAAGCTTACCAATGTTCTTATTAAGAGCTCTCGAGAGAGATGTTGCTCATCAC 473
QY      141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrrArgTyrglyLys 160
DB      474 ACAACGAGAGACTTCCGTAAGCTCATATCACTCTGCTTATGATGATGAAGA 533

```

QY 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
|||  
|||  
DB 534 GATGAGGTCAACTGACCTGGCAAAAATGAGCAAAAATGCTGCAATGAGAAATTTCA 593  
|||  
|||  
QY 181 AspLysHisThrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200  
|||  
|||  
DB 594 AACAAAGCTTATATATGATGACACTTATCATGAGATTTTGGCCACAGAGACAGACAG 653  
|||  
|||  
QY 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGlnGluIleLeuLysSerLeu 220  
|||  
|||  
DB 654 ATTAATGCTACTTGTGATCACTACCAAGATGCAATTTGACAGATATCAACAAGACCTG 713  
|||  
|||  
QY 221 GlnGlnGluLysAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
|||  
|||  
DB 714 ---AGGCTATCCAAAGACGAGTTCTCTGTCTTAAAGACACACTGGAAGTCTTG 770  
|||  
|||  
QY 241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260  
|||  
|||  
DB 771 ATCGTCCCGAAGATGACTTCTGAGAGAGTTGCTTGCGCATCAACAGCGAGAAC 830  
|||  
|||  
QY 261 AspGlnGlyAlaLeuThrArgIleValIleThrArgAlaGluIleAspLeuLysValIle 280  
|||  
|||  
DB 831 GATGAAGAGACCCCTTACAGAGTGTGCGCACAGGCGTGAAGTTGTAAGAAACATA 890  
|||  
|||  
QY 281 GlyGlnGluTyrGlnArgArgAsnSerIleProLeuGlnLysAlaIleThrLysAspThr 300  
|||  
|||  
DB 891 GCAGAGTGAATATCAAAAGAGAGAGAGTGTCTCTTGACGTGCTATTTAAGACACACC 950  
|||  
|||  
QY 301 ArgGlyAspTyrGlnLysMetLeuValAlaLeuLeuGlyGlnLysAspAspAla 317  
|||  
|||  
DB 951 ACCGGTGAATATGAGAAATGCTTGTGGCTTTAAGACATGATGATGCT 1001  
|||  
|||

## RESULT 6

US-10-424-599-1190  
; Sequence 1190, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovacic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; TITLE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 1190  
; LENGTH: 1393  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_101073C.1  
US-10-424-599-1190

## Alignment Scores:

Prod. No.: 4.08e-118 Length: 1393  
Score: 1177.50 Matches: 226  
Percent Similarity: 84.54% Conservative: 42  
Best Local Similarity: 71.29% Mismatches: 48  
Query Match: 73.09% Indels: 1  
DB: 18 Gaps: 1

US-10-690-564-2 (1-317) x US-10-424-599-1190 (1-1393)

QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGlnGlnLeu 20  
|||  
|||  
DB 133 ATGGCGACCTTAAAGTTCTCAGCTCTTCCCTCCCTTCAAGATGATTTGTGAGCGACTC 192  
|||  
|||  
QY 21 ArgThrAlaPheGlnGlyTTPGlyThrAsnGluAspLeuIleLeuSerIleLeuAlaHis 40  
|||  
|||  
DB 193 CGAAAGCCCTTCACAGTTGGGGAACCTAACAGAGAGCTGATTTGATCGATCTTGCTCAT 252  
|||  
|||

QY 41 ArgSerAlaGlnGlnArgLysValIleArgGlnAlaTyrHisGluThrTyrGlyGluAsp 60  
|||  
|||  
DB 253 AGGAATGCTGCTCAGAGAACTTATCAAGAACTTAATGATCCCACTTATGAGAAAGAT 312  
|||  
|||  
QY 61 LeuLeuLysThrLeuAspLysGluLeuSerAsnAspPheGluValArgAlaIleLeuLeuTyr 80  
|||  
|||  
DB 313 CTCTCAAGCCCTTGAACAAAGACTCAGAGTGAATTTTGAGAGCTGGTCTCATCTTTGG 372  
|||  
|||  
QY 81 ThrLeuGluProGlyGlnArgAspAlaLeuLeuAlaAsnGlnAlaThrLysArgTyrThr 100  
|||  
|||  
DB 373 ACACCTGATTCGTGCTGACAGTGAATGATCTTTTGGCAACAGAGCGACTAAATAAGACT 432  
|||  
|||  
QY 101 SerSerAsnGlnValLeuMetGlnValAlaCysThrArgThrSerThrGlnLeuLeuHis 120  
|||  
|||  
DB 433 TCAGGAATCAAGTTCTGTGTGAAATAGCTGCATAGGTCTGTAACAAATGTTTGCT 492  
|||  
|||  
QY 121 AlaArgGlnAlaTyrHisAlaArgTyrLysLysSerLeuGlnGluAspValAlaHisHis 140  
|||  
|||  
DB 493 GCAGAGAGGCTTACCATGTTCTTTATAGAAAGTCTTGAGAGAGATTTGCTCATCAC 552  
|||  
|||  
QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGlnGly 160  
|||  
|||  
DB 553 ACAACAGAGAGACTTCGTAAGCTCATACCTGCTAGTTAGTTATCATGATGAAGGA 612  
|||  
|||  
QY 161 AspGluValAsnMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180  
|||  
|||  
DB 613 GATGAGTCAACTTGAACCTTGCAGAAACCTGAGGCGCAAAATGTCGATAGAAATTTCA 672  
|||  
|||  
QY 181 AspLysHisThrAsnAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200  
|||  
|||  
DB 673 AACAAAGCTTATATATATGACGCTTCATCAGAGATTTTGGCCACAGAGACAGACAG 732  
|||  
|||  
QY 201 IleAsnAlaThrPheAsnArgTyrGlnAspAspHisGlyGlnGluIleLeuLysSerLeu 220  
|||  
|||  
DB 733 ATTAATGCTACTTGTGAATCATCAACAAAGATGCTTTGGAAGATATCAACAGACCTG 792  
|||  
|||  
QY 221 GlnGlnGlyAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240  
|||  
|||  
DB 793 ---AGGCTGATCCAAAGAGAGAGTGTCTCTTGAAGACACTGTGAAGTCTTG 849  
|||  
|||  
QY 241 ThrArgProGluLeuTyrPheValAspValLeuArgSerAlaIleAsnLysThrGlyThr 260  
|||  
|||  
DB 850 ATCCGTCGAGAGAGTACTTCAGAAAGTTGTTGCTTGCGATTCACAGAGAGAAC 909  
|||  
|||  
QY 261 AspGlnGlyAlaLeuThrArgIleValIleThrArgAlaGluIleAspLeuLysValIle 280  
|||  
|||  
DB 910 GATGAAGAGACCCCTTACAGAGTGTGCGCACAGGCGTGAAGTTGTAAGAAACATA 969  
|||  
|||  
QY 281 GlyGlnGluTyrGlnArgArgAsnSerIleProLeuGlnLysAlaIleThrLysAspThr 300  
|||  
|||  
DB 970 GCAGATGATATCAAAAGAGAGAGAGTGTCTCTTGAACGTGCTATTTAAGACACACC 1029  
|||  
|||  
QY 301 ArgGlyAspTyrGlnLysMetLeuValAlaLeuLeuGlyGlnLysAspAspAla 317  
|||  
|||  
DB 1030 ACCGGTGAATATGAGAAATGCTTGTGGCTTTAAGACATGATGATGCT 1080  
|||  
|||

## RESULT 7

US-10-219-220-17  
; Sequence 17, Application US/10219220  
; Publication No. US20030082724A1  
; GENERAL INFORMATION:  
; APPLICANT: Flihm, Barry  
; APPLICANT: Lasham, Annette  
; TITLE OF INVENTION: Compositions affecting programmed cell  
; FILE OF INVENTION: death and their use in the modification of plant development  
; CURRENT APPLICATION NUMBER: US/10/219,220  
; FILING DATE: 2002-08-14  
; PRIOR APPLICATION NUMBER: U.S. No. US20030082724A1 09/325,932  
; FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 290  
; SOFTWARE: FastSeq for Windows Version 3.0



```
; SEQ ID NO 17
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-219-220-17

Alignment Scores:
Pred. No.: 1,66e-117 Length: 1293
Score: 1171.50 Matches: 226
Percent Similarity: 83.60% Conservative: 39
Best Local Similarity: 71.29% Mismatches: 51
Query Match: 72.72% Indels: 1
DB: Gaps: 1

US-10-690-564-2 (1-317) x US-10-219-220-17 (1-1293)

QY 1 MetAlaThrLeuLyValSerAspSerValProAlaProSerAspAlaGluGlnLeu 20
   1171.50
   83.60%
   71.29%
   72.72%

DB 95 ATGTGCACTCTCACCGCCGCGACCACTGCCCTGTAGCCGATGATCGACGACGCTC 154

QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
   155 CGGACAGCCTTCGCGAGATGGGAAACAATGAGAAGCTGATCATATCCATATTGGGTCAT 214

QY 41 ArgSerAlaGluGlnArgLyValIleArgGlnAlaTyriHisGluThrTyrgLyGluAsp 60
   215 AGGAATGGCGCGCAGAGGAAGCTGATTCCGCAAACTATGCCAGACTTACGCGCAGGAC 274

DB 61 LeuLeuLyThrLeuAspLyGluLeuSerAspAspPheGluArgAlaIleLeuLeuTrp 80
   275 CTCCTCAAGGCATTTGACAGAGAACTTACCAATGATTTCCAGAGGCTGGTGGCTTTGG 334

QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGlnAlaThrLyAspTrpThr 100
   335 TCACCTTGATCCGCGTGAACGTGATGCTACTTGGCGAATACGACGAAAGAATGGACT 394

DB 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuHis 120
   395 TCAGAGCAACCAAGGTTCATGAGAAATAGCTTGACAGAGGTCTCGGACAGGTTGGTTAG 454

QY 121 AlaArgGlnAlaTyriHisAlaArgTyriLySerLeuGluGluAspValAlaHisHis 140
   455 GCAAGACCAACATATCATGCCGATACAGAAAGTCAATGAAAGGACGTCGCTCACAC 514

DB 141 ThrThrGlyAspPheArgLyLeuLeuValSerLeuValThrSerTyraArgTyrgLyGlu 160
   515 ACAACTGGAATTTTGTGAAGTTGCTGGTACTCTTGGGAGCTCTACCGTAATAGGA 574

QY 161 AspGluValAsnMetThrLeuAlaLyGlnGluAlaLySerValHisGluLySilelys 180
   575 GATGAGGTGAATATGACTTTGGCAAAACGAGGCTAAGTACTCCACGGAAGATCTCA 634

DB 181 AspLyHisTyriAsnAspGluAspValIleArgIleLeuSerThrArgSerLySilaGln 200
   635 GAGAAGGCTTATGCGCCATGAGATCTCATAGGATTTTGGCTACTAGAGCAAAAGCAG 694

QY 201 ILAspAlaThrPheAsnArgTyriGlnAspAspHisGlyGluGluIleLeuLySerLeu 220
   695 GTCAATGCTACGCGTAATCTCACTAATGAGTTGGAAATGATATCAACAGATTGG 754

DB 221 GluGluGlyAspAspAspAspLyPheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
   755 AAAAAT---GATCAAAAGAAGCGGTTCTTACTATCTAGAGCTACAGTAAAGTGGCT 811

QY 241 ThrArgProGluLeuTyriPheValAspValLeuArgSerAlaIleAsnLyThrGlyThr 260
   812 ACTCGCCCTGAGAGATTTTGAAGGTTCTTGTCTACCATCAATTAAGCAAGAGAAC 871

DB 261 AspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLyValIle 280
   872 GATGAAGGGCTCTGACCAAGATAGTTGCTACAGAGGCGCAGGTTGACGAAAGTTTAA 931

QY 281 GlyGluGlyTyriGlnArgAsnSerIleProLeuGluLySilaIleThrLyAspThr 300
```

```
DB 932 AGTGAAGATACACAGAGGAGATAGCATCCCTCTGCATCGGCCATTGTCAAGACACT 991

QY 301 ArgGlyAspTyriGluLyMetLeuValAlaLeuLeuGlyGluAspAla 317
   992 ACTGAGACTATGAAAAAATGCTTCTGCGATGATGGCAGCTGAGGCT 1042

RESULT 8
US-10-393-840-109
; Sequence 109, Application US/10393840
; Publication No. US20030229922A1
; GENERAL INFORMATION:
; APPLICANT: Blockberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1012c3
; CURRENT APPLICATION NUMBER: US/10/393,840
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: PaetSeq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Pinus radiata
US-10-393-840-109

Alignment Scores:
Pred. No.: 1,66e-117 Length: 1293
Score: 1171.50 Matches: 226
Percent Similarity: 83.60% Conservative: 39
Best Local Similarity: 71.29% Mismatches: 51
Query Match: 72.72% Indels: 1
DB: Gaps: 1

US-10-690-564-2 (1-317) x US-10-393-840-109 (1-1293)

QY 1 MetAlaThrLeuLyValSerAspSerValProAlaProSerAspAlaGluGlnLeu 20
   95 ATGTGCACTCTCACCGCCGCGACCACTGCCCTGTAGCCGATGATCGACGACGCTC 154

DB 95 ATGTGCACTCTCACCGCCGCGACCACTGCCCTGTAGCCGATGATCGACGACGCTC 154

QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
   155 CGGACAGCCTTCGCGAGATGGGAAACAATGAGAAGCTGATCATATCCATATTGGGTCAT 214

DB 155 CGGACAGCCTTCGCGAGATGGGAAACAATGAGAAGCTGATCATATCCATATTGGGTCAT 214

QY 41 ArgSerAlaGluGlnArgLyValIleArgGlnAlaTyriHisGluThrTyrgLyGluAsp 60
   215 AGGAATGGCGCGCAGAGGAAGCTGATTCCGCAAACTATGCCAGACTTACGCGCAGGAC 274

DB 215 AGGAATGGCGCGCAGAGGAAGCTGATTCCGCAAACTATGCCAGACTTACGCGCAGGAC 274

QY 61 LeuLeuLyThrLeuAspLyGluLeuSerAspAspPheGluArgAlaIleLeuLeuTrp 80
   275 CTCCTCAAGGCATTTGACAGAGAACTTACCAATGATTTCCAGAGGCTGGTGGCTTTGG 334

DB 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGlnAlaThrLyAspTrpThr 100
   335 TCACCTTGATCCGCGTGAACGTGATGCTACTTGGCGAATACGACGAAAGAATGGACT 394

QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuHis 120
   395 TCAGAGCAACCAAGGTTCATGAGAAATAGCTTGACAGAGGTCTCCGACAGGTTGGTTAG 454

DB 121 AlaArgGlnAlaTyriHisAlaArgTyriLySerLeuGluGluAspValAlaHisHis 140
   455 GCAAGACCAACATATCATGCCGATACAGAAAGTCAATGAAAGGACGTCGCTCACAC 514

QY 141 ThrThrGlyAspPheArgLyLeuLeuValSerLeuValThrSerTyraArgTyrgLyGlu 160
```

```
|||||
Db 515 ACAACGTGAGATTTCTGTAAGTCTGCTACTCTTGGAGACTCTTACCGTAAGATGGA 574
Qy 161 AEPGLVALAENMETHTHLEUALALYSGINGUALALYSEUVALHISGLUTYSILEYS 180
Db 575 GATAGAGTGAATATGACTTTGGCAAAAGCAGAGCTTAAGATCTCCAGAGAGATCTCA 634
Qy 181 ASPLYSHTSTYRANAPBGLUAPVALILEARGILELEUSERHTHARGSERLYSALGLN 200
Db 635 GAGAAAGCTTATGGCCATGAGATCTCATAGAGATTTGGCTACTAGAGCAAGCAG 694
Qy 201 ILEASNALATHRPHAEANARGTYRGINAPBAPHISGLYGLUGLUILEULYSERLEU 220
Db 695 GTCAATGCTACGCTGAATCACTACAAAATAGATTGGAAATGATATCAACAGAGATTG 754
Qy 221 GLUGLUGLYAPBAPAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPB 240
Db 755 AAAACT---GATCCAAAGAGCGCTTCTTACTATCTAGAGAGCTTACGTAAGTGCCTG 811
Qy 241 THTRARGPROGLULEUTYRPHAEVALAPBVALLEUARGSERLALILEANLYSTHGLYTH 260
Db 812 ACTGCGCTGAGAGATTTTGAAGAGTCTTGTCTAGCCATCAATTAAGCGAGAGAA 871
Qy 261 AEPGLUGLYALALEUTHRARGILEVALTHTRARGALAGLUILEAPBAPBAPBAPBAPB 280
Db 872 GATBAGAGGGCTCTGACCAAGATGTTGCTACCGAGCGGAGTTGACATGAGATTATA 931
Qy 281 GLYGLUGLYTARGINARGASERILEPROLEUGLULYSALAILERTHLYSAPBTH 300
Db 932 AGTGAAGAGTACCAAGAGAGATAGCATCCCTCTCGATCGTGCATTTGTCAAGAGACT 991
Qy 301 ARGGLYASPTYRGILULYSMETLEUVALALEULEUGLYGLUAPBAPBAPBAPB 317
Db 992 ACTGAGACATATGAAAAAATGCTTGTGCAATGATGCGCACTGCGAGGCT 1042
```

## RESULT 9

```
US-10-219-220-218
; Sequence 218, Application US/10219220
; Publication No. US20030082724A1
; GENERAL INFORMATION:
; APPLICANT: Film, Barry
; APPLICANT: Laeham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE OF INVENTION: death and their use in the modification of plant development
; FILE REFERENCE: 11000.1022cl
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US/10/219, 220
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 290
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 218
; LENGTH: 1058
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-219-220-218
```

## Alignment Scores:

```
Pred. No.: 2.26e-113 Length: 1058
Score: 1132.50 Matches: 218
Percent Similarity: 83.91% Conservative: 48
Best Local Similarity: 68.77% Mismatches: 50
Query Match: 70.30% Indels: 1
Gaps: 1
```

US-10-690-564-2 (1-317) x US-10-219-220-218 (1-1058)

```
Qy 1 METALATHREULYSVALSERAPSERVALPROALPROSERAPBAPBAPBAPBAPBAPB 20
Db 72 ATGCGGCTATCGGGTGCACCGCTCGTCCGCTCGGCTGAGAGTGCAGAGACTC 131
Qy 21 ARGTHRALAPHEGLUGLYTPGLYTHASGLUAPBAPBAPBAPBAPBAPBAPBAPBAPB 40
```

```
Db 132 CAAAAGCTTTCGAGATGCGGAGACGAATGAAGATCTGATTCATTCACACTGCGCTCAC 191
Qy 41 ARGSERLALUGLULARGYSVALILEARGLINALTYRHSGLUTHTYRGLYGLUAPB 60
Db 192 AGAAGCGACGCGAGCGGAAAGTAATCCGCAAAACATATATCCGACATATGCGGAAAGAT 251
Qy 61 LEULEULYETHRLEUAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPB 80
Db 252 CTCTCTCAAGCGCTTGAAGAGAACTCTTATGAGCTTGTAGAGATCTGTGCTGTGTG 311
Qy 81 THTRLEUGLUPROGLULARGAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPB 100
Db 312 ACCCTGATCTCTCGAGAGCTGATGCACTTCTTCCAAATGACCTTCCAAAGATGACT 371
Qy 101 SERSERANGLIVALLEUWETGLUALALCYETHRARGTHRSERHTHGLULEULHIS 120
Db 372 TCGAGCAACTGCGGCTTCAATGAAATGCTTGCACAGAGGCTTCAATGAGATTATTCATG 431
Qy 121 ALARGLALATHRYTHISALARGTYRGLYSERLEUGLULAPBAPBAPBAPBAPBAPB 140
Db 432 GTAGGCGAGGCTATATGATGCTGTATTAAGAAATCTTGAAAGACATCGCATATCAC 491
Qy 141 THTRGLYAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPB 160
Db 492 ACTACGCGGATTTCCGCAAGCTGCTTCTTCTCGCAAGTACTTTCGTTATGAGGG 551
Qy 161 AEPGLVALAENMETHTHLEUALALYSGINGUALALYSEUVALHISGLUTYSILEYS 180
Db 552 CTTGAGGTGAACATGACATGCGAGATGACAGAGCTTAAGATCTCATGAGAGATTAC 611
Qy 181 ASPLYSHTSTYRANAPBGLUAPVALILEARGILELEUSERHTHARGSERLYSALGLN 200
Db 612 GAGAAAGCTTACATATGATAGATGCTCATCAGAAATGTTATCAAGAAAGTAAAGCTCAG 671
Qy 201 ILEASNALATHRPHAEANARGTYRGINAPBAPHISGLYGLUGLUILEULYSERLEU 220
Db 672 CTATATGCAACCTCAATTAATCAACATGAGTTTGGAAATGCCATCAACAGATCTG 731
Qy 221 GLUGLUGLYAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPBAPB 240
Db 732 ---AAGCGTATCCAAATGATGAAATTTCTGAACCTCGTGAATCAACAGCTGAGATCTG 788
Qy 241 THTRARGPROGLULEUTYRPHAEVALAPBVALLEUARGSERLALILEANLYSTHGLYTH 260
Db 789 ACTTATCCAGAGAGATCTTGAAGAGTACCTGCGTCCGCTCAACAGCTGAGAA 848
Qy 261 AEPGLUGLYALALEUTHRARGILEVALTHTRARGALAGLUILEAPBAPBAPBAPBAPB 280
Db 849 GATGAATGGGCTTTTACAGAGTACCACTCGTGGAGAGTTGACATGACGCGGATC 908
Qy 281 GLYGLUGLYTARGINARGASERILEPROLEUGLULYSALAILERTHLYSAPBTH 300
Db 909 AAGAACAGTACCAAGAAAGAAAGAGTGTCTCTGATCGGCTATTCGAATGACACA 968
Qy 301 ARGGLYASPTYRGILULYSMETLEUVALALEULEUGLYGLUAPBAPBAPBAPB 317
Db 969 TCTGGGAGCTTACAAAGAAAGATGCTTCTGCTTGTGATGACATGAGATGCT 1019
```

## RESULT 10

```
US-10-425-114-10772
; Sequence 10772, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
```

```
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 10772
/ LENGTH: 1145
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: 700970311_FLI
US-10-425-114-10772

Alignment Scores:
Pred. No.: 3,83e-111 Length: 1145
Score: 1112.50 Matches: 225
Percent Similarity: 83.70% Conservative: 42
Best Local Similarity: 70.53% Mismatches: 49
Query Match: 69.06% Indels: 5
DB: 18 Gaps: 1

US-10-690-564-2 (1-317) x US-10-425-114-10772 (1-1145)

QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20
   |||||
DB 54 ATGGCGACCTTAAGGTTCTCAGCCTCTTCCCGCTTGCAAGATGATGTGACAGCTC 113
   |||||

QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
   |||||
DB 114 CGAAAAGCCTTCCAGGTTGGGAACTAACGAGAGCTGATTGATGATCTTGCTCAT 173
   |||||

QY 41 ArgSerAlaGluGluLysValIleArgGlnAlaTyrHisGluThrTyrGlyLysAsp 60
   |||||
DB 174 AGGAATCTGCTCAGAGGAGCTTAATCAGAAACTTATCCAGACCTATGAGAAAGAT 233
   |||||

QY 61 LeuLeuLysThrLeuAspLysGluLeuSerAspAspPheGluLysArgAlaIleLeuLeuTyr 80
   |||||
DB 234 CTCTCTCAAGGCTTGACCAAGAACTCAAGATGATTGTAGAGGCTGTTCAATCTTTGG 293
   |||||

QY 81 ThrLeuGluProGlyLysArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrThr 100
   |||||
DB 294 ACAGTTGATCTCTGCTGAGCTGATGATTTTGGCAACGAGCGACTAAATAATGAGCT 353
   |||||

QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGluLeuHis 120
   |||||
DB 354 TCAGAGAAATCAGGTTCTGGGAAATAGCCTGCATGAGGCTCTGAAACAATGTTGGCT 413
   |||||

QY 121 AlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluLysAspValAlaHisHis 140
   |||||
DB 414 GCGAGAAAGCTTAACCACTGTTCTTTAAGAGTCTCTGAGAGAGATGTTGCTCATAC 473
   |||||

QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160
   |||||
DB 474 ACAACAGGAGACTTCCGTAAAGCTCATACTCTGTTAGTCTTATGCAATATGAGGA 533
   |||||

QY 161 AspGluValAspMetThrLeuAlaLysGlnGluAlaLysLeuValHisGluLysIleLys 180
   |||||
DB 534 GATGAGGTCAACTTGACCTT-GCAAAAACCTGACCAAAATTCGTGCATGAGAAATTGCA 592
   |||||

QY 181 AspLysHisTyrAspAspGluAspValIleArgIleLeuSerThrArgSerLysAlaGln 200
   |||||
DB 593 AACAAAGCTTAATATGATGAGACTTCATGAGATTTTGGCCACAAAGAGACAGACAA 652
   |||||

QY 201 IleAsnAlaThrPhe-Asn-ArgTyrGlnAspAspHisGlyGluGluIleLeuLysSerI 220
   |||||
DB 653 ATTAAGCTACTTTTGAATCCACTACAAAGATGATTTGGAACAAGATATCAACAAGGA-C 711
   |||||

QY 220 euGluGluGlyAspAspAspLysPheLeuAlaLeuLeuArgSerThrIleGlnCysL 240
   |||||
DB 712 TG---AAGGCTGATCCAAAGACAGAGTTCCTGTCTTACTTAAGACAACTGTAAGTGTCT 768
   |||||

QY 240 euThrArgProGluLeuTyrPheValAspValLeuLysSerAlaIleAsnLysThrGlyT 260
   |||||
DB 769 TGATCCGTCCCGAAGATCTTGAAGAGGTATTTGCTGGGATCAACAAGCGAGGA 828
   |||||

QY 260 hrAspGluGlyAlaLeuThrArgIleValThrThrArgAlaGluIleAspLeuLysValI 280
   |||||

DB 829 CCGATGAAGAGACCCCTTAACAAGTGTGTGCCCAAGGGCTAGGTTGATTTGAAAAACA 888
   |||||

QY 280 legGlyGluGlyTyrGlnArgArgAsnSerIleProLeuGluLysAlaIleThrLysAspT 300
   |||||
DB 889 TAGCAGATGATATCAAGAGAGACAGTGTCTCTTGAAGCTCTATTTGTAAGACA 948
   |||||

QY 300 hrArgGlyAspTyrGluLysMetLeuValAlaLeuLeuGlyGluLysAspAla 317
   |||||
DB 949 CCAACGGTACTATGAGAAATGCTTGTGGCCCTTTAGAGACATGATATGCT 1001
   |||||

RESULT 11
US-10-424-599-70457
/ Sequence 70457, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21 (53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 70457
/ LENGTH: 1321
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_34633C.1
US-10-424-599-70457

Alignment Scores:
Pred. No.: 7.41e-104 Length: 1321
Score: 1046.50 Matches: 213
Percent Similarity: 81.56% Conservative: 48
Best Local Similarity: 66.56% Mismatches: 55
Query Match: 64.96% Indels: 4
DB: 18 Gaps: 1

US-10-690-564-2 (1-317) x US-10-424-599-70457 (1-1321)

QY 1 MetAlaThrLeuLysValSerAspSerValProAlaProSerAspAspAlaGluGluLeu 20
   |||||
DB 90 ATGGCGACCTTAAGGTTCTCAGCCTCTTCCCTTCCATGGAAAGACGTAGCAATTTG 149
   |||||

QY 21 ArgThrAlaPheGluGlyTyrGlyThrAsnGluAspLeuLeileSerIleLeuAlaHis 40
   |||||
DB 150 CGAAAAGCTTTTCAAGGGTGGGAAAGCAATGAAGATGATTAATCAATCTGGGTTGAC 209
   |||||

QY 41 ArgSerAlaGluGluLysValIleArgGlnAlaTyrHisGluThrTyrGlyLysAsp 60
   |||||
DB 210 ACAAAATGCGGCTCAGCGTAAAGTTGATCTGTGAAGCTTATAGTACAAACCATGAGAAAGAT 269
   |||||

QY 61 LeuLeuLysThrLeuAspLysGluLeuSerAspAspPheGluLysArgAlaIleLeuLeuTyr 80
   |||||
DB 270 CTCTTAAGATTTTGAACAAGAACTTTCAAGTGAATTTGACGGGCTGTGTGGTTGG 329
   |||||

QY 81 ThrLeuGluProGlyLysArgAspAlaLeuLeuAlaAsnGluAlaThrLysArgTyrThr 100
   |||||
DB 330 ACATTTGATCTCTGCTGAGGTTGATGCTTTTATGATTAAGAACCACTAAGATGTTGACT 389
   |||||

QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuHis 120
   |||||
DB 390 TCAAAACAATGGGTATTTTGAAGAAATGCTTCACTAGATCTTCACTTGATCTCTTAAG 449
   |||||

QY 121 AlaArgGlnAlaTyrHisAlaArgTyrLysSerLeuGluGluLysValAlaHisHis 140
   |||||
DB 450 GCAAGCAAGCAATATCAAGCCCGTTTCAAAAAGTCCCTTGAAGAAAGACCTTATCAT 509
   |||||

QY 141 ThrThrGlyAspPheArgLysLeuLeuValSerLeuValThrSerTyrArgTyrGluGly 160
   |||||
```



```

/ FEATURE:
/ OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS5887_1
US-10-767-701-13525

Alignment Scores:
Pred. No.: 5,21e-100 Length: 1385
Score: 1011.50 Matches: 192
Percent Similarity: 77.78% Conservative: 53
Best Local Similarity: 60.95% Mismatches: 69
Query Match: 62.79% Indels: 1
DB: 19 Gaps: 1

US-10-690-564-2 (1-317) x US-10-767-701-13525 (1-1385)

QY 1 MetAlaThrLeuIleuValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
DB 58 ATGGCGAGCTCAAGGCTCCAGCGACGCTCCCTCCGCTCCGATGACGCGAGCGCTC 117
QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
DB 118 CGCAGAGCATTCACAGGGTGGGCGACGACGAGCGCTCATCTGCTCCATCTCGGCAC 177
QY 41 ArgSerAlaGluGlnArgValIleArgGlnAlaTrpHisGluThrTyrglyGluAsp 60
DB 178 CGCGACGCCCGCAGCGCCGCGCATCCGCGCCCTACCGCCGACGCGCCAGCGAGAG 237
QY 61 LeuLeuThrLeuAspLeuValSerAspAspPheGluArgAlaIleLeuLeuTrp 80
DB 238 CTCTCTCGCTCCATCACGAGACGATCTCCGCGACTTCAGAGGGCGGTGATCTGTGG 297
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIleValTrpThr 100
DB 298 ACGCTGGACCCAGCGGAGCGCGACGCGGTCTGGCCACAGAGCGCGCGAGAGTGCGAG 357
QY 101 SerSerAsnGlnValLeuMetGluValAlaCysThrArgThrSerThrGlnLeuHis 120
DB 358 CCAAGGAACCGCGTCTGCTGAGATCGCTGACGCGCACTTCGCGCGAGGTCTTCGCG 417
QY 121 AlaArgGlnAlaTrpHisAlaArgTrpValSerLeuGluGlnAspValAlaHisHis 140
DB 418 GCCAGCGAGCGGTATCCAGCGGCTTACGCGCTTACGAGGAGCATCGCGCCAC 477
QY 141 ThrThrGlyAspPheArgIleLeuLeuValSerLeuValThrSerTrpArgTrpGlyGlu 160
DB 478 GTCACTGGGAGCTTCGCAAGCTACTGTCGCACTTGAATGATACGCTATGATGGA 537
QY 161 AspGluValaIleMetThrLeuAlaValGlnGluAlaValLeuValHisGluValIleVal 180
DB 538 CCAGAGGTCAACCAAGGTGGCACACTCTAGAACCCAACTACTCCATGAGAGATCCAT 597
QY 181 AspIleHisIleTrpAsnAspGluAspValIleArgIleLeuSerThrArgSerIleValGln 200
DB 598 CACAAGGCTTACAGTGAATGATGATATCATGAACTTCAACACTAGAGACCAACTCAG 657
QY 201 IleAsnAlaThrPheAsnArgTrpGlnAspAspHisGlyGluGlnIleLeuValSerLeu 220
DB 658 CTGCTGCTCATTCATCATTCATTAACAATGATGCAATTTGGCCACCGAATCAACAAAGATCTG 717
QY 221 GluGluGlyAspAspAspAspIlePheLeuAlaLeuLeuArgSerThrIleGlnCysLeu 240
DB 718 --AAGGCTACCCCAAGATGAGTACTCAAAACACTGGGGCAATCATCCGCTGCTTC 774
QY 241 ThrArgProGluLeuTrpPheValAspValLeuArgSerAlaIleAsnIleValThrGlyTrp 260
DB 775 AGTTGCCCCCAKAGTCTTTGAGAAGGTGCGCAGGAGGCAATAGCGGGGCTTAGGACA 834
QY 261 AspGluValaIleLeuThrArgIleValIleThrArgAlaGluIleAspLeuValIle 280
DB 835 GACGAGAAAGCCCTGACGAGGCTCATTAACACCGCGCTGAGGTGAGCTGAAACGATG 894
QY 281 GlyGluGlyTrpGlnArgArgAsnSerIleProLeuGlnIleValAlaIleThrIleValAspTrp 300
DB 895 AAGGAGCGTACCAAGAGAAACAGCGTCCCTGAGCGCGCTGCGTGGGAGACCC 954

```

```

QY 301 ArgGlyAspTrpGlyMetLeuValAlaLeuLeuGlyGluAsp 315
DB 955 TCCGGCGACTACGAGACGATGCTCTCGCGCTCTTGGCAGGAG 999

RESULT 14
US-10-425-114-21986
/ Sequence 21986, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaka, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 21986
/ LENGTH: 1235
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3356-011-F3_FLI
US-10-425-114-21986

```

```

Alignment Scores:
Pred. No.: 7.34e-100 Length: 1235
Score: 1009.50 Matches: 191
Percent Similarity: 77.46% Conservative: 53
Best Local Similarity: 60.63% Mismatches: 70
Query Match: 62.66% Indels: 1
DB: 18 Gaps: 1

US-10-690-564-2 (1-317) x US-10-425-114-21986 (1-1235)

QY 1 MetAlaThrLeuIleuValSerAspSerValProAlaProSerAspAspAlaGluGlnLeu 20
DB 78 ATGGCAACCTCAAGGCTCCAGCGACTGTCTCCGCTCCGATGACGCGACGAGCTG 137
QY 21 ArgThrAlaPheGluGlyTrpGlyThrAsnGluAspLeuIleIleSerIleLeuAlaHis 40
DB 138 CGAAGCAATTCAGAGGAGGAGGACGAGCGCTATCATCTCCATTCCTCGGCCAC 197
QY 41 ArgSerAlaGluGlnArgValIleArgGlnAlaTrpHisGluThrTyrglyGluAsp 60
DB 198 CGGAGCGCGCGAGCGCGCGCCATCCGCGCGCTTACCGCGAGGCTTACGCGAGAG 257
QY 61 LeuLeuThrLeuAspLeuValSerLeuSerAspAspPheGluArgAlaIleLeuLeuTrp 80
DB 258 CTCTCTCGCTCATTCAGCGAGCATCTCCGCGACTTCGAGAGGCGCTGATCTCTGCG 317
QY 81 ThrLeuGluProGlyGluArgAspAlaLeuLeuAlaAsnGluAlaThrIleValAspTrp 100
DB 318 AGCTTGACCCAGCGGAGCGGAGCGGTGCTGGCCGAGAGCGCGCGGAGAGTGAGAG 377
QY 101 SerSerAsnGlnValLeuMetGluValaIaCysThrArgThrSerThrGlnLeuHis 120
DB 378 CCGCGGACCGGTGTGCTGCGAGATCGCTTCAAGCGGCACTCCGCGGAGATCTTCGCG 437
QY 121 AlaArgGlnAlaTrpHisAlaArgTrpValSerLeuGlnGluAspValaIleHis 140
DB 438 ACCAGGACGCTTATCAACAGCGGTTCAGAGCCTCGCTTGAAGAGACATCTCGCGCAC 497
QY 141 ThrThrGlyAspPheArgIleLeuValSerLeuValIleThrSerTrpArgTrpGlyGlu 160
DB 498 GTCACTGGGAGCTTCGCAAGGCTATTTGGCCACTTGAAATGATTAACGCTATGATGGA 557
QY 161 AspGluValaIleMetThrLeuAlaValGlnGluAlaValLeuValHisGluValIleVal 180

```



**This Page Blank (uspto)**